Kasper Boye and Petar Kehayov (Eds.)
Complementizer Semantics in European Languages

Empirical Approaches to Language Typology

Editors
Georg Bossong
Bernard Comrie
Kristine Hildebrandt

Volume 57

Complementizer Semantics in European Languages

Edited by
Kasper Boye and Petar Kehayov

DE GRUYTER
MOUTON

ISBN 978-3-11-041651-0
e-ISBN (PDF) 978-3-11-041661-9
e-ISBN (EPUB) 978-3-11-041666-4
ISSN 0933-761X

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

© 2016 Walter de Gruyter GmbH, Berlin/Boston
Typesetting: PTP-Berlin, Protago-TEX-Production GmbH, Berlin
Printing and binding: CPI books GmbH, Leck
♾ Printed on acid-free paper
Printed in Germany

www.degruyter.com

Foreword

The idea for this book arose in connection with the Workshop on "Semantic functions of complementizers in European languages", which we organized in October 28–29, 2011, at the University of Copenhagen, Denmark. Around two thirds of the book chapters are elaborations on contributions to this workshop, the remaining one third arose independently of the workshop.

Our work on the book has benefitted from the help and support from many people and several institutions. We would like to thank the contributors and our editor with De Gruyter Mouton, Julie Miess, for their patience with us and their continued interest in the project. Thanks are also due to the now closed LANCHART Centre at the University of Copenhagen and its leader Frans Gregersen for sponsoring the above-mentioned workshop, to the Department of Nordic Studies and Linguistics at the University of Copenhagen for granting Kasper Boye research assistance in connection with his work on the volume, and to the Alexander von Humboldt Foundation, the Department of Slavic Philology at the University of Regensburg and the Graduate School for East and Southeast European Studies at the University of Regensburg and Ludwig Maximilian University of Munich for the research assistance provided to Petar Kehayov. Finally, we are grateful to Marie-Louise Lind Sørensen, Jessie Leigh Nielsen and Roman Fisun for excellent editorial assistance.

Kasper Boye & Petar Kehayov, Copenhagen & Regensburg 2016

Table of contents

Foreword —— v

Petar Kehayov and Kasper Boye
Complementizer semantics – an introduction —— 1

Part I: **Indo-European languages**

Peter McQuillan
Finite clause complementizers in Celtic —— 15

Benjamin Fagard, Paola Pietrandrea and Julie Glikman
Syntactic and semantic aspects of Romance complementizers —— 75

Jackie Nordström and Kasper Boye
Complementizer semantics in the Germanic languages —— 131

Björn Hansen, Alexander Letuchiy and Izabela Błaszczyk
Complementizers in Slavonic (Russian, Polish, and Bulgarian) —— 175

Axel Holvoet
Semantic functions of complementizers in Baltic —— 225

Brian D. Joseph
The semantics and syntax of complementation markers as an areal phenomenon in the Balkans, with special attention to Albanian —— 265

Natalia Serdobolskaya
Semantics of complementation in Ossetic —— 293

Yaron Matras and Anton Tenser
Complementizers in Romani —— 341

Part II: **Non-Indo-European languages**

Xabier Artiagoitia and Arantzazu Elordieta
On the semantic function and selection of Basque finite complementizers — 379

Albert Borg and Ray Fabri
Semantic functions of complementizers in Maltese — 413

Petar Kehayov
Complementation marker semantics in Finnic (Estonian, Finnish, Karelian) — 449

Kristina Kotcheva and Michael Rießler
Clausal complementation in Kildin, Skolt and North Saami — 499

Gerson Klumpp
Semantic functions of complementizers in Permic languages — 529

Erika Körtvély
Modal functions of the complementizer *hogy* 'that' in Hungarian — 587

Birsel Karakoç and Annette Herkenrath
Clausal complementation in Turkish and Noghay in a semantic perspective — 619

Mikhail Knyazev
Complementizers in Kalmyk — 665

Natalia Serdobolskaya
Semantics of complementation in Adyghe — 691

Diana Forker
Complementizers in Hinuq — 745

Peter Harder
Complementizers and the complementizing function – structure and pragmatics in a panchronic perspective — 793

Petar Kehayov and Kasper Boye
Complementizer semantics in European languages: Overview and generalizations —— 809

Language index —— 879

Subject index —— 885

List of contributors

Xabier Artiagoitia
University of the Basque Country UPV/EHU
E-Mail: xabier.artiagoitia@ehu.eus

Izabela Błaszczyk
University of Regensburg
E-Mail: Izabela.Blaszczyk@sprachlit.uni-regensburg.de

Albert Borg
University of Malta
E-Mail: albert.borg@um.edu.mt

Kasper Boye
University of Copenhagen
E-Mail: boye@hum.ku.dk

Arantzazu Elordieta
University of the Basque Country UPV/EHU
E-Mail: arantzazu.elordieta@ehu.eus

Ray Fabri
University of Malta
E-Mail: ray.fabri@um.edu.mt

Benjamin Fagard
Lattice, CNRS, ENS, PSL Research University, University of Paris 3
E-Mail: benjamin.fagard@ens.fr

Diana Forker
University of Bamberg
E-Mail: diana.forker@uni-bamberg.de

Julie Glikman
University of Strasbourg
E-Mail: julieglikman@yahoo.fr

Björn Hansen
University of Regensburg
E-Mail: Bjoern.Hansen@sprachlit.uni-regensburg.de

Peter Harder
University of Copenhagen
E-Mail: harder@hum.ku.dk

Annette Herkenrath
Justus Liebig University Giessen
E-Mail: annette.herkenrath@germanistik.uni-giessen.de

Axel Holvoet
University of Vilnius
E-Mail: axel.holvoet@flf.vu.lt

Brian Joseph
Ohio State University
E-Mail: bjoseph@ling.ohio-state.edu

Birsel Karakoç
Uppsala University
E-Mail: birsel.karakoc@lingfil.uu.se

Petar Kehayov
University of Regensburg, University of Tartu
E-Mail: petar.kehayov@ur.de

Gerson Klumpp
University of Tartu
E-Mail: klumpp@ut.ee

Mikhail Knyazev
Saint Petersburg State University
E-Mail: misha.knjazev@gmail.com

Kristina Kocheva
University of Freiburg
E-Mail: kristina.kotcheva@skandinavistik.uni-freiburg.de

Erika Körtvely
University of Regensburg
E-Mail: kortvely@t-online.de

Alexander Letuchiy
Higher School of Economics, Moscow
E-Mail: aletuchiy@hse.ru

Yaron Matras
University of Manchester
E-Mail: yaron.matras@manchester.ac.uk

Peter McQuillan
University of Notre Dame du Lac
E-Mail: Peter.T.McQuillan.4@nd.edu

Jackie Nordström
Goethe University of Frankfurt
E-Mail: nordstroem@em.uni-frankfurt.de

Paola Pietrandrea
LLL CNRS, University of Tours
E-Mail: pietrand@uniroma3.it

Michael Rießler
University of Freiburg
E-Mail: michael.riessler@skandinavistik.uni-freiburg.de

Natalia Serdobolskaya
Russian State University for the Humanities, Moscow State Pedagogical University
E-Mail: serdobolskaya@gmail.com

Anton Tenser
Vanderbilt University, University of Helsinki
E-Mail: toxa7@hotmail.com

List of gloss abbreviations

1	first person
2	second person
3	third person
9	noun class 9
A	agent-like argument of canonical transitive verb
A/DA	Finnic first infinitive
ABE	abessive
ABL	ablative
ABS	absolutive
ABST	abstract
ACC	accusative
ACNMLZ	action nominalization
ACT	active voice
AD	location 'near'
ADD	additive particle
ADE	adessive
ADJ	adjective marker
ADV	adverb(ial) marker
ADVS	adversative particle
AFF	affirmative
AFFC	affective
AGR	agreement
ALL	allative
ALOC	'animate' location (case)
ANT	anterior
ANTIP	antipassive
AOR	aorist
ART	article
ASSOC	associative
AT	location 'at, by'
ATR	free attributive marker
ATTR	attributive
AUG	(morphonological) augment
AUT	autonomous (verb form)
AUTQ	autoquotative particle
AUX	auxiliary
BEN	benefactive
BOR	borrowing
C	clause
CAUS	causative
CIT	citative
CL	clitic
CLF	classifier
CMP	comparative
CN	connegative
CNTRF	counterfactive
COM	comitative
COMP	complementizer or complementation marker
COMPL	completive

COND	conditional
CONJ	conjunction
CONT	location with contact
COP	copula
COR	correlative pronoun
CORE	non-nominative core article
CS	complementation strategy
CSE	case
CTP	complement taking predicate
CVB	converb
CVBANT	anterior converb
CVBI	imperfective converb
CVBP	participle-converb
CVBTEMP	temporal converb
D	gender agreement marker
DAT	dative
DEF	definite
DEI	deixis
DEM	demonstrative
DEP	dependent
DER	derivational affix
DET	determiner
DIM	diminutive
DIR	directive
DIRC	directional case
DMS	dental modal subordinator
DO	direct object
DUR	durative
DYN	dynamicity
EGR	egressive case
ELA	elative case
EMPH	emphatic
EPIST	epistemic
EQU	equative
ERG	ergative
EVID	evidential
EXP	experiencer
EXST	existential copula
EXT	extension
F	feminine
FCT	factivity
FIN	finite
FRQ	frequentative aktionsart
FUT	future
GEN	genitive
GEN1	first genitive
GEN2	second genitive
GER	gerund
GNT	general tense
HAB	habitual
HORT	hortative

HPL	human plural
I	gender agreement marker
IDEO	ideophone
II	gender agreement marker
III	gender agreement marker
IJ	interjection
ILL	illative
ILOC	'inanimate' location
IMP	imperative
IMPS	impersonal
IN	inessive/illative
IND	indicative
INDF	indefinite
INDIR	indirective
INE	inessive
INF	infinitive
INFM	infinitive marker
INS	instrumental
INT	interjection
INTF	intensifier
INTFUT	intentional future
INTR	intransitive marker
IO	indirect object or agreement with indirect object
IPFV	imperfective
IQ	indirect question
IRR	irrealis
ITER	iterative
IV	gender agreement marker
J	gender agreement marker
JUSS	jussive
LAT	lative
LOC	locative
LOCUT	locutor person marker
LOG	logophoric
M	masculine
MA	Finnic third infinitive
MAL	malefactive
MOD	modal
MSD	masdar
N	neuter
NACT	nonactive
NARR	narrative
NEC	necessitive
NEG	negation
NMLZ	nominalizer or nominalization
NOM	nominative
NOMACT	nomen actionis
NONLOCUT	non-locutor person marker
NPI	negative polarity item
O	objective conjugation
OBJ	object

OBL	oblique or (in Adyghe) a valency-changing prefix that promotes an oblique to argument position
OBM	oblique (evidential) mood
ONO	onomatopoeia
OPT	optative
P	person
PAG	agent participle
PART	partitive
PASS	passive
PFV	perfective
PL	plural
PNEC	necessive participle
POSS	possessive
POST	postterminality
POSTP	person agreement on postpositions
POT	potentialis
POVB	postverb
PP	perfect participle
PPA	past participle active
PPP	past participle passive
PQ	polar question marker
PR	person agreement with the possessor
PRED	predicative
PREF	prefix
PREP	preposition
PRET	preterite
PRF	perfect
PRO	pronoun
PROG	progressive
PROH	prohibitive
PROX	proximal
PRP	present participle
PRPA	present participle active
PRS	present
PRT	particle
PRVB	preverb
PST	past
PTCP	participle
PURP	purposive
Q	question marker
QUOT	quotative
R	realis
RE	refactive/reversive
REC	recent past tense
REFL	reflexive
REFL/MID	reflexive-middle
REL	relativizer or relative pronoun
REM	remote past
RES	resultative
ROOT	root
SBJ	subject

SBJV	subjunctive
SG	SG
SIM	similative word
SML	simultaneous
SPESPL	superessive-superlative
SPL	superlative
SPR	location ‘on’
SUB	(adverbial) subordinator
SUBLAT	sublative
TEMP	temporal
TOP	topic
TRANS	translative
UNC	uncertainty
UWPST	unwitnessed past
V	gender agreement marker
VN	verbal noun
VOC	vocative
WH	content question marker

Petar Kehayov and Kasper Boye

Complementizer semantics – an introduction

1 Background

Complementizers may be defined as conjunctions that have the function of identifying clauses as complements (e.g. Crystal 2008: 93; Noonan 2007: 55). In recent years, it has become increasingly clear that complementizers may have additional functions. Some of these additional functions may be characterized as pragmatic. For instance, optional complementizers like English *that* may serve as pause markers (e.g. Kaltenböck 2009) or as signals of genre or social level. Other functions are semantic in the sense that they represent conventional contributions to the meanings of the complements. The present book is concerned with semantic complementizer functions.

The recognition of such functions owes a lot to Frajzyngier (e.g. Frajzyngier & Jasperson 1991; Frajzyngier 1995), who demonstrated that complementizers can have modal functions (see also Nordström 2010; Boye, van Lier & Brink 2015; Wiemer & Letuchiy, in progress). For instance, complementizers frequently indicate degree of certainty about the proposition expressed by the complement. In Jacaltec (Mayan), the complementizer *chubil* indicates "a high degree of credibility or certainty", whereas *tato* "introduces a notion of disbelief or reservation about a hearsay" (Craig 1977: 267); thus, (1) expresses a higher degree of certainty than (2) about the proposition 'the president is going to come'.

Jacaltec

(1) *Xal* *naj* *alcal* [***chubil*** *chuluj* *naj* *presidente*].
said CLF/the alcalde **COMP** will.come CLF/the president
'The alcalde said that the president is going to come'.
(Craig 1977: 268)

(2) *Xal* *naj* [***tato*** *chuluj* *naj* *presidente*].
said CLF/he **COMP** will.come CLF/the president
'He said that the president is going to come'.
(Craig 1977: 268)

In Lango (Nilo-Saharan), likewise, the complementizer *ká* indicates uncertainty about the complement proposition, and contrasts with the epistemically neutral complementizer *nî*.

Lango

(3) *Dákô* *párô* *àpárâ* [***ká*** *ɲákô* *òrègò* *kál*.]
woman 3SG.consider.HAB consider.GER **comp** girl 3SG.grind.PFV millet
'The woman doubts whether the girl ground the millet'.
(Noonan 1992: 227)

(4) *ɲákô* *òkòbbì* *dákô* [***nî*** *dyèl* *òcàmò*].
girl 3SG.say.BEN.PFV woman **COMP** goat 3SG.eat.PFV
'The girl told the woman that the goat ate it.'
(Noonan 1992: 220)

The semantic functions of complementizers are not restricted to the domain of modality, however. In Awa Pit (Barbacoan), the complementizers *ka* and *sa* not only indicate uncertainty about the complement proposition, but also location in time relative to the matrix clause: *ka* indicates uncertainty about non-future propositions (5), and *sa* about future propositions (6) (Curnow 1997: 256–259). The two complementizers contrast with epistemically neutral absence of a complementizer (7).

Awa Pit

(5) *Na=na* *min-tu-s* [*us=na* *kwa-t* ***ka***].
1SG.(NOM)=TOP think-IPFV-LOCUT 3SG.(NOM)=TOP eat.PFV.PTCP **COMP**
'I am wondering if he has eaten'.
(Curnow 1997: 257)

(6) *Na=na* *Marcos=ta=na* *mɨma-ta-w* [*mizhaka=ma*
1SG.(NOM)=TOP Marcos=ACC=TOP ask-PST-LOCUT:SBJ when=INTER
a-mtu ***sa***].
come-IPFV.PTCP **COMP**
'I asked Marcos when he would come'.
(Curnow 1997: 257)

(7) *Gregorio=na* [*ashaŋpa=na* *az-tu* ***Ø***] *kizh-tɨ-zi*.
Gregorio=TOP woman=TOP cry-IPFV.PTCP say-PST-NONLOCUT
'Gregorio said that his wife was crying'.
(Curnow 1997: 260)

In a number of languages, complementizers serve to distinguish propositional (truth-valued) complements from state-of-affairs (non-truth-valued) complements (Ransom 1986; Boye 2012: 206–213; see also Dik & Hengeveld 1991 and Boye 2010 specifically on perception-predicate complements). In Latvian (Indo-European), the complementizer *kā* marks state-of-affairs complements, whereas *ka* marks propositional complements.

Latvian

(8) *Viņš* ***zināja,*** [***kā*** *spēlēt* *klavieres*].
he **know.PST.3** **COMP** play.INF piano.ACC[PL]
'He knew how to play the piano'.
(Axel Holvoet, this volume, and p.c.)

(9) *Viņš* ***zināja,*** [***ka*** *viņa* *spēlē* *klavieres*].
he **know.PST.3** **COMP** she play.PRS.3 piano.ACC[PL]
'He knew that she played the piano'.
(Axel Holvoet, this volume, and p.c.)

In Tukang Besi (Austronesian), absence of complementizer marks state-of-affairs complements, whereas the complementizer *kua* marks propositional complements.

Tukang Besi

(10) *No-'ita-'e* [***Ø*** *no-kanalako* *te* *osimpu*]
3R-see-3OBJ 3R-steal CORE young.coconut
'She saw him stealing the coconut'.
(Donohue 1999: 403)

(11) *No-'ita-'e* [***kua*** *no-kanalako* *te* *osimpu*].
3R-see-3OBJ **COMP** 3R-steal CORE young.coconut.
'She saw that he had stolen the coconut'.
(Donohue 1999: 404)

2 Research questions

Complementizer semantics is still heavily understudied. Accordingly, the following questions remain unanswered.

A. How common is it for complementizers to have semantic functions?
   - Frajzyngier (1995) and Nordström (2010) claim that all complementizers are primarily modal, but Boye, van Lier & Brink (2015) argue that complementizers need not have semantic functions at all.

B. Which kinds of semantic functions may complementizers have?
   - They can be modal and temporal, and they may distinguish state-of-affairs from propositions, but can they also be, for instance, aspectual, or concerned with polarity.

C. Which kinds of semantic distinctions are found in complementizer systems?
   - Complementizer systems (i.e. distributionally delimited sets of complementizers) may code binary semantic contrasts (for instance, between certainty and uncertainty). But are more complex semantic contrasts – and thus, larger systems – also found?

D. How do semantic complementizer functions develop diachronically?
   - Complementizers are known to typically develop out of demonstrative pronouns, adverbs or prepositions, but how do (some of them) come to express e.g. degree of certainty or location in time?

Answers to these questions are not only interesting in themselves. They may cast new light on the following topics:

a. Complementizer optionality
   - Research on complementizer optionality (e.g. Elsness 1984; Kaltenböck 2009; Shank & Cuyckens 2010) focuses on factors governing the "omission" of complementizers. The question of what qualifies a complementizer for optionality in the first place remains unanswered. The qualification may be, at least partly, functional (Boye, van Lier & Brink 2015).

b. Complementizer combination
   - In some languages, two or more distinct complementizers may under certain circumstances be combined in the same clause. A detailed description of semantic complementizer functions may provide insights into such combinations, including restrictions on the relative order in which complementizers may occur.

c. Subordinator functions
   - In some languages, the selfsame expression serves both as a complementizer and as an adverbial subordinator and/or relativizer (cf. e.g. English *if* and *that*). A better understanding of semantic complementizer func-

tions may thus pave the way for a better understanding of the functions of relativizers and adverbial subordinators.

d. Modality and other notional categories
   – Complementizers are largely ignored in research on notional categories such as modality and time. A detailed study of complementizers and their semantic functions may, however, turn out to reveal insights that are crucial for our understanding of those categories.

3 Focus and limitations of this book

In attempting to answer the research questions given in A–D above, European languages are a good place to start. In a crosslinguistic study based on a genealogically stratified sample of the world's languages, Boye, van Lier & Brink (2015: 6) found that 43 out of 76 languages have complementizers (in the sense defined by Noonan [2007: 55]), but in European languages the ratio of languages with complementizers appears to be even higher. Moreover, European languages often have systems of complementizers in which individual complementizers contrast functionally, with clear semantic effects of substituting one for another.

This book therefore brings together specialists of all language families in Europe to describe semantic complementizer functions as well as those phenomena on which semantic complementizer functions may be assumed to have a direct descriptive or explanatory impact: complementizer systems, diachrony of complementizers, complementizer omission, and complementizer combination.

3.1 Languages

Table 1 presents the languages discussed in this book, together with their genetic affiliation. The languages presented in brackets play only a marginal role in the individual contributions and as a rule give background information as to the principal languages discussed. As can be seen from the table, all indigenous language families of the European continent are covered.

Table 1: Languages with their genetic affiliation

phylum	languages
Indo-European	
Celtic	Irish, Welsh
Romance	Portuguese, Spanish, French, Italian, Sardinian, Romanian (Aromanian)
Germanic	Icelandic, Faroese, Norwegian, Swedish, Danish, German, Yiddish, Dutch, Afrikaans, Frisian, English
Slavic	Russian, Polish, Bulgarian (Macedonian)
Baltic	Latvian, Lithuanian
Greek	Greek
Albanian	Albanian
Indo-Iranian	Ossetic, Romani
Basque	Basque
Afro-Asiatic	Maltese
Uralic	
Finnic	Estonian, Finnish, Karelian, Ludian
Sami	North Saami, Kildin Saami, Skolt Saami
Permian	Komi-Zyrian, Komi-Permyak, Udmurt
Ugric	Hungarian
Turkic	Turkish, Nogai
Mongolic	Kalmyk
West Caucasian	Adyghe
East Caucasian	
Tsezic	Hinukh (Khvarshi, Tsez)
Avar-Andic	(Avar, Bagvalal, Ghodoberi)
Dargi	(Dargwa)
Lezgic	(Archi, Lezgi, Tsakhur)
Nakh	(Bats, Chechen, Ingush)

3.2 Focus

This book does not cover all aspects of complementizers, nor does it include everything that has been dealt with in the literature under the term *complementizer*.

Firstly, the focus is on complementizer *semantics*. This entails that we do not intend to cover for instance the morphosyntax of complementizers in all details. But obviously, the book contains a lot of information also on morphosyntax as distributional facts, co-occurrence restrictions, etc. are crucial for understanding the semantics of complementizers.

Secondly, the focus is on complementizers in a narrow sense. Some scholars use the term *complementizer* in a wide sense, including complement markers

side by side with conjunctions in adverbial and relative clauses, and even relative adverbs and pronouns (this is especially the case in different variants of Generative Grammar and frameworks such as Head-driven Phrase Structure Grammar, inspired by Generative Grammar, where complementizers and complementizer phrases play an important role). This book focuses on complementizers in the sense of complementation markers. In many cases, however, complementizer forms have additional subordinating functions as well, as in the case of English *that*, which can be used also to mark relative and adverbial clauses. Moreover, it may be difficult to tear complementizers apart from other subordinators (see Section 3.4 below), and the synchronic and diachronic relations between complementizers and other subordinators may provide important clues to complementizer semantics. Accordingly, many chapters include detailed discussions of phenomena included under a broad understanding of complementizers.

Thirdly, the book focuses on complementizers in balanced (including finite) complements and the closest equivalents of such complementizers (see Stassen 1985 on the distinction between *balancing* and *deranking*). One reason for this is that complementizers are more easily or safely identified in balanced than in deranked (including nonfinite) complements: the features that identify complements as deranked may themselves be complementation or subordination markers, which makes the role of any element under suspicion for being a complementizer more dubious (cf. Section 3.4 below). Another reason is our intention to ensure as high a degree of comparability as possible. Obviously, there are limitations to the degree to which complementizers and complementizer systems can be compared. For one thing, not all languages have balanced complementation. For such languages, as mentioned, the focus is on the closest functional counterparts of complementizers found in balanced complements. Still, many of the chapters in this book include detailed information also on deranked complementation, as this may be crucial to the description of complementizer contrasts and thus complementizer semantics.

3.3 Recurrent topics

In order to ensure a high degree of comparability, the contributors dealing with individual language families were asked to address the following issues, if possible.

The inventory of complementizers.

- Which complementizers can be identified in the languages covered?
- How can they be identified?

The semantics of complementizers
- What are the semantic (conventional) functions of the complementizers dealt with, in addition to marking complements?
- What are the arguments for the semantic analyses (including, possibly, minimal pairs, cooccurrence restrictions, harmonic combinations, distribution)?

The distribution of complementizers
- In which types of complements are the complementizers found?
- With which types of complement-taking elements are the complementizers found?
- Do the complementizers under consideration make up a "system", i.e. a distributionally delimited set of expressions? If yes, how many members does the system have?
- Are there elements (e.g. negation or modal verbs) that trigger one specific complementizer?

Complementizer optionality
- Are any of the complementizers optional (similarly to English *that* in *I know (that) the butler did it*)? If yes, which complementizers, under which circumstances, and with which effect?

Combinability issues
- Can any of the complementizers be combined with other complementizers? If yes, which complementizers, under which circumstances, and with which effect?
- Can any of the complementizers be combined with other subordination markers (relativizers, adverbializers)? If yes, which complementizers, which other subordination markers, under which circumstances, and with which effect?

Non-complementizing functions of complementizer forms
- Can any of the complementizers be used as relativizers (similarly to English *that* in *the book that I read two days ago*) or adverbializers (similarly to English *that* in *He is so tired that he can hardly keep his eyes open*)? If yes, which complementizers, and under which circumstances?
- Can any of the complementizers be used as e.g. adverbs, particles, verbs or nouns? If yes, which complementizers, under which circumstances, and with which effect?

Diachrony
- What are the diachronic sources of the complementizers, and how did they develop?
- How did the semantic functions of the complementizers under investigation develop?

While more or less all of the contributions on individual language families deal with more or less all of these issues, there is considerable variation in how and to what extent the issues are (and can be) dealt with.

3.4 The identification of complementizers

Dealing with the issues introduced in the previous section obviously presupposes that complementizers – defined, as mentioned, in accordance with Noonan (2007: 55) – can be identified in the first place. For different reasons, however, the identification of complementizers is not always a straightforward matter (see Boye, van Lier & Brink 2015 for discussion), and in some cases there may be as good reasons for identifying a candidate for a complementizer as something else than a complementizer as there are for identifying it as a complementizer.

A first set of problems was touched upon in Section 3.2 above. It concerns the association of what we can refer to as complementizing function – i.e. the function of identifying complements – with a given expression (the suspected complementizer). To establish such an association is much less troublesome for balanced than for deranked dependent clauses. Balanced dependent clauses are by definition identical to independent clauses. Thus, in balanced complements which have an extra element compared to an independent clause, this element is an obvious candidate for a complementizer. In English *that* complements, for instance, *that* is what distinguishes the complement (e.g. *that she married him*) from an independent clause (*she married him*). In the case of deranked dependent clauses, things are more complicated. It may be impossible to isolate one complementizing element from a cluster of elements (e.g. nominalizers, particles, word order, case-marking on arguments etc.) that are together responsible for the deranked status of the complement.

A second and more profound set of problems have to do with the identification of complementizing function in the first place, Expressions that may be suspected of having complementizing function often have a number of other functions as well. For instance, English *that* serves also to mark relative clauses, adverbial clauses and (insubordinated) independent clauses. This means that on a monosemic approach, *that* could not be said to have a complementizing func-

tion in the strict sense which is relevant here. Rather, one would have to posit a more general function that subsumes the complementizing function (even "subordinating function" might not be general enough). This raises the question whether it is at all possible to identify a complementizing function: does the more general function encompass complementizing function or does it replace it and make it superfluous. This question is not less difficult to answer for cases where suspected complementizer forms also serve as modal particles or even verbs.

Even in cases where a suspected complementizer form does not have other functions than identifying complements – or where, at least, it makes sense to ignore such functions – identifying a complementizing function may present problems. One reason for this is that the identification depends on a clear definition of what a complement is. Thus, for instance, if you consider purposive clauses as complements you may identify a marker of such clauses as a complementizer. But if you do not consider them complements, it would be odd to identify a complementizing function.

Problems in identifying complementizers – as well as possible solutions – are discussed in several of the contributions to this book.

4 General structure

After this introductory chapter come the chapters on individual language families all of which deal with the semantic functions complementizers may have in addition to their complementizing function. These chapters are followed by a conclusive chapter in which, based on the contributions to this book, we give an overview of the data from European languages and dare some generalizations.

References

Boye, Kasper. 2012. *Epistemic meaning. A crosslinguistic and functional-cognitive study* (Empirical Approaches to Language Typology 43). Berlin: Mouton de Gruyter.

Boye, Kasper. 2010. Reference and clausal perception-verb complements. *Linguistics* 48(2). 391–430.

Boye, Kasper, Eva van Lier & Eva T. Brink. 2015. Epistemic complementizers: A crosslinguistic survey. *Language Sciences* 51(5). 1–17.

Craig, Colette. 1977. *The structure of Jacaltec*. Austin: University of Texas Press.

Crystal, David. 2008. *A dictionary of linguistics and phonetics*, 2nd edn. Oxford: Blackwell.

Curnow, Timothy J. 1997. *A grammar of Awa Pit (Cuaiquer): An indigenous language of southwestern Colombia*. Canberra: Australian National University dissertation.

Dik, Simon C. & Kees Hengeveld. 1991. The hierarchical structure of the clause and the typology of perception-verb complements. *Linguistics* 29. 231–259.
Donohue, Mark 1999. *A grammar of Tukang Besi*. Berlin: Mouton de Gruyter.
Elsness, Johan. 1984. *That* or zero: A look at the choice of object clause connective in a corpus of American English. *English Studies* 65(6). 519–533.
Frajzyngier, Zygmunt. 1995. A functional theory of complementizers. In Joan Bybee & Suzanne Fleischman (eds.), *Modality in grammar and discourse*, 473–502. Amsterdam: John Benjamins.
Frajzyngier, Zygmunt & Robert Jasperson. 1991. That-clauses and other complements. *Lingua* 83. 133–153.
Kaltenböck, Gunther. 2009. Is *that* a filler? On complementizer use in spoken object clauses. *Vienna English Working Paper* 18(1). 28–63.
Noonan, Michael. 1992. *A grammar of Lango*. Berlin: Mouton de Gruyter.
Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description 2*: *Complex constructions*, 2nd edn. 52–150. Cambridge: Cambridge University Press.
Nordström, Jackie 2010. *Modality and subordinators*. Amsterdam: John Benjamins.
Ransom, Evelyn N. 1986. *Complementation: Its meaning and forms*. Amsterdam: John Benjamins.
Shank, Christopher & Hubert Cuyckens. 2010. Moving beyond synchrony: Applying a diachronic corpus-based multivariate analysis to examine the development and use of *I + think* as an 'epistemic parenthetical'. Paper presented at Societas Linguistica Europaea 2010, Vilnius, 1–5 September.
Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.
Wiemer, Björn & Alexander Letuchiy (eds.). In progress. Clausal complements in Eurasian languages. [Special issue]. *Linguistics*.

Part I: **Indo-European languages**

Peter McQuillan

Finite clause complementizers in Celtic

1 Introduction

1.1 Complementation and complementizers

Noonan (1985: 42) describes *complementation* as "the syntactic situation that arises when a notional sentence or predication is an argument of a predicate". Furthermore, a *complementizer* (1985: 44–45) is defined as "a word, particle, clitic, or affix whose function it is to identify the entity as a complement". The focus of this paper will be on Celtic finite complementizers, basing the analysis on Irish and Welsh. Despite fundamental differences between the two languages, the complementizer systems of both show a basic affinity. This lies in the fact that both languages make a fundamental distinction between complement clauses that are epistemically neutral on the one hand, and epistemically uncertain on the other as to the realization of their propositions, a distinction consistently reflected at the formal level. These I will call Type 1 and Type 2 respectively (borrowing the terminology of Nordström & Boye, this volume).[1] The Type 1 complementizers are of disparate origin: Irish *go* comes from a preposition meaning 'to, as far as', while Welsh *y* is originally an affirmative preverbal particle occurring both in main and subordinate clauses. In the case of Type 2, both languages have a preverbal polar question marker and simply transpose a direct polar question as a complement clause to a main predicate. A noteworthy feature of Celtic complementizers is that they have separate negative forms historically unrelated to the affirmative ones. Irish further distinguishes itself from Welsh in that it has separate past tense forms for all its complementizers.

1.2 Organization of the discussion

This section introduces the forms of the complementizers (Section 1.3) to be discussed subsequently. Section 2 introduces characteristic features of the Celtic lan-

1 I am borrowing here the terminology used by Kasper Boye and Jackie Nordström in their presentation 'The semantics of finite clause complementizers in Germanic languages' given as part of the LANCHART workshop on Semantic Functions of Complementizers in European Languages, University of Copenhagen, 29 October 2011. See further this volume.

guages as they affect our discussion of complementation, while Section 3 deals with the semantics of the two types of complement clauses in Irish and Welsh. Initially, the semantics of complement types will be investigated through the examination of minimal pairs where the sole difference between each member of the pair lies in the choice of complementizer. Section 4 then looks at the distribution of each type across main predicate types, especially positive propositional attitude, knowledge and utterance predicates for Type 1 and predicates of asking and uncertainty for Type 2. The issue of the possibility or otherwise of complementizer omission is dealt with in Section 5. Section 6 examines the extent to which complementizers may combine with other subordinators specifically in adverbial clauses, while Section 7 discusses the non-complementizing functions of complementizers and complement markers, in relative and adverbial clauses for instance, as well as in polar questions. Finally, Section 8 looks at the diachronic origins and development of complementizers in both languages.

1.3 Celtic finite complement clauses and their complementizers

Finite complements in Celtic must always occur after their governing main clause predicates. Complementizers are always clause initial; the possibility of omitting the complementizer varies between the two languages: as we shall see in Section 5 below, Irish and Welsh differ somewhat in this respect. We can divide complementizers according to the two semantic types introduced in Section 1.1 above; in addition, each type can be subdivided into non-focus and focus forms: non-focus forms are followed directly by the verb of the complement, focus forms by a non-verbal constituent of the complement which has been fronted for reasons of contrast and emphasis.

1.3.1 Type 1 non-focus affirmative

Irish affirmative complements of this type are introduced by the particle *go*. This causes an initial mutation (see Section 2.2 below) to a following consonant or vowel known as *nasalization* (or *eclipsis*). There is a separate past tense form *gur* which causes no mutation.

Welsh is syntactically more complex (see especially Section 2.5 and Section 4.1.1.2 below); see also King (1993: 302–10) and Borsley et al. (2007: 75–85). Affirmative complements of this type are introduced by:

1. the affirmative particle *y* when the verb of the complement is future tense or conditional mood.
2. the verbal noun *bod* ('be') when the verb of the complement is present, imperfect or perfect tense. Each of these three tenses is periphrastic in nature and employs *bod* and its inflected forms as an auxiliary. In this case, complement marking is incorporated in the clause initial verb form and the main verb of the complement appears as a verbal noun.
3. the preposition *i* 'in' when the proposition of the complement is anterior to that of the main clause. The preposition is followed by the complement subject as its indirect object; this is in turn followed by a complement verb in the form of its verbal noun. This allows 3. to be an alternative to 2. above in certain cases, particularly where the perfect tense is involved.

Section 2.5 below will feature a brief characterization of the Welsh verbal system. Both 2. and 3. while formally non-finite are regarded as syntactically finite by Welsh grammarians. The arguments for this position will be presented in the relevant sections below.

1.3.2 Type 1 non-focus negative

Negative complementizers in both Irish and Welsh are distinguished by the fact that they are etymologically distinct from their affirmative counterparts:

1. In Irish the form is *nach* (followed by nasalization of the verb, see Section 2.2); there is a separate past tense form *nár* which causes no mutation.
2. In Welsh the form is *na* which causes the initial mutation called lenition (Section 2.2) to the following verb; the form before a vowel is *nad*.

Welsh negative complements are syntactically simpler than their affirmative counterparts in that in all cases *na* is followed by the relevant inflected forms of the verb no matter which tense or mood. So in negative complements the syntactic distinctions in Section 1.3.1 above are neutralized.

1.3.3 Type 1 focus

Both Irish and Welsh have distinctive forms introducing a focus complement clause. Since the normal word order of Celtic clauses is Verb-Subject-Object (see Section 2.1), a focus complement means a complement beginning with a fronted element, which is not a finite verb.

1. Irish distinguishes formally in the verb 'be' between a substantive verb *tá* and a copula *is* (see Section 2.4); if the complement position is followed immediately by a non-verbal, focused element), then the indirect speech forms of the copula are used to introduce the clause: affirmative *gur* in the present (*gurb* before a vowel); *gurbh* in the past and conditional; negative *nach* in the present, in the past tense *nár*, with *nárbh* before a vowel.
2. In Welsh the corresponding affirmative form is the invariant *mai*, in origin a relative form of the verb 'be'; the negative form is *nad*.

1.3.4 Type 2 non-focus

Type 2 complementizers in Celtic are in origin polar interrogative markers. Wh-question markers are not a focus of this discussion with one exception: Irish *cé acu* 'which of them' which functions like a Type 2 complementizer and is equivalent to English *whether...or not* in complements.

In Irish the form of the polar interrogative is *an* (nasalizing) for all non-past forms; the form in the past is *ar*.

Once again, distribution in Welsh is more involved and calls for some comment (King 1993: 302–310). In particular, the system is reminiscent of Type 1 in that the present and imperfect forms of the verb *bod* are treated differently from other verb forms.

1. The polar interrogative marker *a* corresponds to Type 1 *y* (Section 1.3.1 above) in that it introduces a complement containing a future or conditional verb form; it causes lenition to a following consonant (see Section 2.2 below).
2. The present tense of the verb 'be' has a paradigm of polar question forms inflected for person and number (e.g., 3SG *ydy*, 3PL *ydyn*). These forms essentially replace Type 1 *bod* as complementizer (Section 1.3.1. 2. above) in Type 2 clauses in that they introduce complements containing the periphrastic forms of present, imperfect and perfect.

See Section 2.5 for more on Welsh verb categories.

1.3.5 Type 2 focus

In both Irish and Welsh, focus complementizers for Type 2 have different forms from Type 1; as is the case for non-focus complementizers, they are also polar interrogatives:

1. In Irish the indirect question forms of the copula are used to introduce the clause: affirmative *ar* in the present (*arb* before a vowel); *arbh* in the past and conditional.
2. In Welsh the corresponding affirmative form is the invariant *ai*.

1.4 Overview of complementizers

The following are the complementizers in summary form.[2] Note: in Irish, all non-past particles cause nasalization (or eclipsis), this is indicated by the superscript ^{N}; in the case of the past tense markers, the verb is typically lenited in simple sentences and remains so in complementation. In Welsh, *y* causes no mutation (in more formal registers the form is *yr* before a vowel or before *h*), while Type 2 *a* and Type 1 negative *na* cause lenition, likewise indicated by a superscript ^{L}.

Table 1: Complementizer forms when directly followed by verb

IRISH TYPE 1		**IRISH TYPE 2**	
AFFIRMATIVE		**AFFIRMATIVE**	
Non-past	***go*^{N}**	**Non-Past**	***an*^{N}**
Past	***gur***	**Past**	***ar***
			cé acu
NEGATIVE			
Non-past	***nach*^{N}**		
Past	***nár***		
WELSH TYPE 1		**WELSH TYPE 2**	
AFFIRMATIVE	***y***	**AFFIRMATIVE**	***a*^{L}**
	bod/bo (+VN)		
	i (+VN)		
NEGATIVE	***na*^{L}*/nad***		

2 For a summary view of Welsh complementation see King (1993: 302–310), Borsley et al. (2007, 75–85); for Irish see Ó Siadhail 1989: 252).

Table 2: Complementizer forms when directly followed by non-verbal element (focus)

IRISH TYPE 1		**IRISH TYPE 2**	
AFFIRMATIVE		**AFFIRMATIVE**	
Non-past	***gur/gurb***	**Non-Past**	***ar/arb***
Past	***gur/gurbh***	**Past**	***ar/arbh***
NEGATIVE			
Non-past	***nach***		
Past	***nár/nárbh***		
WELSH TYPE 1		**WELSH TYPE 2**	
AFFIRMATIVE	***mai***	**AFFIRMATIVE**	***ai***
NEGATIVE	***nad***		

2 The Celtic languages

The modern Celtic languages comprise two subgroups: a Gaelic or Goidelic branch consisting of the Irish, Scottish and Manx (Isle of Man) varieties and a British or Brittonic grouping composed of Welsh, Cornish and Breton. In order to distinguish the modern languages from the long extinct mainland European Celtic languages (Continental Celtic: languages such as Gaulish, Lepontic and Celtiberian) both of these sub-branches are often referred to together as Insular Celtic owing to the fact that the homelands of the speakers of five of the six languages are, or have been, the archipelago of Ireland and Britain. Given the fact that the homeland of what is now called Breton has been in North West France for a millennium and a half, this designation is somewhat incongruous but is nonetheless often used. The two modern groupings are also sometimes distinguished to as Q-Celtic and P-Celtic, referring to the respective outcomes of the Indo-European labio-velar *k^w, which splits into a velar in the Goidelic languages and a labial in the Brittonic ones, as in Irish *cé* and Welsh *pwy* 'who' from *k^wei-, for example. (However, this split is not coterminous with *insular*, as it also affected the continental languages).[3]

In typological terms, Fife (1993: 21–24) characterizes four features as strongly diagnostic of the Celtic languages. Three are particularly relevant here (the fourth

3 For an overview of the Celtic languages from both the linguistic and socio-linguistic perspectives, see the various contributions in Ball & Fife (1993).

is the existence of inflected or conjugated prepositions) and will be dealt with in sections 2.1 to 2.3 below.

2.1 Word order

The Celtic languages show an unmarked declarative word order of Verb-Subject-Object, uniquely among the Indo-European family, as exemplified by (1) Irish and (2) Welsh:[4]

(1) *Cheannaigh* *sí* *an* *teach*
PST\buy 3SG.F ART house
'She bought the house'

(2) *Pryn-odd* *hi* *'r* *ty*
buy-PST.3SG 3SG.F ART house
'She bought the house'

Celtic languages are inclined to front non-verbal elements of the clause for expressive emphasis and contrast (focus); as has already been indicated in Section 1.3, this has consequences for the formal marking of complements.

2.2 Initial mutations

The Celtic languages have a system of initial consonantal mutations which has morphophonemic status, that is to say, it participates in grammatical oppositions. Irish has two mutations: lenition and nasalization (often termed *eclipsis* by Irish grammarians), while Welsh has three: lenition (known by Welsh grammarians as *the soft mutation*), nasalization and spirantization. I will illustrate the principles of these initial mutations using the Irish word *carr* 'car' preceded by *a*, the third person possessive particle. This possessive form is invariant for singular (both masculine and feminine) and plural. What distinguishes each case however is the appropriate mutation on the initial *c-* of the noun. With the feminine singular, there is no mutation:

(3) *a* *carr*
POSS.3SG.F car
'her car'

4 Unless otherwise indicated, examples are my own. I have sometimes shortened examples cited from others for convenience.

With the masculine singular the initial *c-* is lenited, producing in this case a spirant (4), while the plural possessive nasalizes, leading in this instance to voicing of the following consonant (5):

(4) *a* *c**h**arr*
POSS.3SG.M car
'his car'

(5) *a* ***g**carr*
POSS.3PL car
'their car'

Lenition in Irish generally results in spirantization; in Welsh, however, a lenited voiceless stop is voiced, so corresponding to (4) above we have for Welsh *car* 'car':

(6) *ei* ***g**ar*
POSS.3SG.M car
'his car'

If we reconsider example (1) above, we see that the clause-initial verb form *cheannaigh* also begins with a spirant *ch-* resulting from lenition. In most Irish verbs, the form of the past tense (here *cheannaigh*) is characterized in relation to the stem of the verb (*ceannaigh*) by lenition of the initial consonant.

2.3 Pre-verbal particles

The Celtic languages are rich in pre-verbal particles indicating illocutionary force (affirmation, negation, interrogation) and subordination. As already noted in Section 1 above for example, both Irish and Welsh have a polar interrogative marker. In Irish the non-past affirmative form is *an,* while *a* is the corresponding Welsh form. Unlike Welsh, Irish has a separate past form *ar* (7); in the Welsh example, lenition is again realized as voicing (8), compare *prynodd in* (2) above:

(7) *Ar* *cheannaigh* *sí* *an* *teach?*
Q PST\buy 3SG.F ART house
'Did she buy the house?'

(8) *A* *bryn-odd* *hi* *'r* *ty?*
Q PST\buy-3SG 3SG.F ART house
'Did she buy the house?'

The Celtic languages also show evidence of affirmative particles (AFF) preceding the finite verb. In our original Irish example (1), the past tense verb is lenited

although there is no particle overtly preceding it. The lenition however reflects the presence historically of a particle *do* as in (9) below. This particle has dropped out in most situations in most dialects, leaving only the following mutation. The common affirmative in Welsh is the particle *fe* or *mi* (10) both of which lenite:

(9) *Do* *cheannaigh* *sí* *an* *teach*
AFF PST\buy 3SG.F ART house
'She bought the house'

(10) *Fe/Mi* *bryn-odd* *hi* *'r* *ty*
AFF PST\buy-3SG 3SG.F ART house
'She bought the house'

However, the Welsh affirmative particles show the same tendency to omission as the Irish in colloquial registers, so that (10) above is typically realized in speech as (11), where only the mutation remains, as in the original Irish example:

(11) *Bryn-odd* *hi* *'r* *ty*
PST\buy-3SG 3SG.F ART house
'She bought the house'

In general, affirmative particles and their effects are more pervasive in Welsh than in Irish.[5] In the case of (8) above, the interrogative *a* is also subject to deletion in speech. Therefore the distinction between declarative and interrogative sentences generally becomes one of intonation (Borsley et al. 2007: 36). This greater tendency towards deletion in Welsh than in Irish leaves its mark on complementation as will be apparent below.

2.4 Verb 'be' in Irish

As indicated in Section 1, Irish has two verbs 'be':

1. the verb *tá* (known as the 'substantive' verb) which typically links the subject to a predicate adjective or adverbial. The principal forms are present *tá*, past *bhí*, future *beidh* and conditional *bheadh*; in addition, both present and past have special dependent forms used after certain particles and conjunctions. The present dependent form is *fuil* (negative *níl* is from *ní fhuil*), the past is *raibh*.

5 See Borsley et al. (2007: 34–37) for an overview of Welsh preverbal particles.

2. the verb *is* (known as the *copula*) which typically links the subject to a predicate noun or pronoun. The principal forms are present *is*, past and conditional *ba* (there is no future tense).

In complement-taking predicates, both 1. and 2. may be used before adjectives likes *cinnte* 'certain', *fíor* 'true' or *ceart* 'right'.

2.5 The Welsh verb

In order to elucidate some of the points made in Section 1 above and to facilitate subsequently more detailed discussion, I present as succinctly as I can an outline (somewhat simplified) of the Welsh finite verb as it pertains to complementation.[6] The critical distinction is between 1. simple and 2. periphrastic categories (distinctions between them are illustrated by the English verb 'buy'):

1. Preterite (*bought*), future (*will buy*) and conditional (*would buy*): these are all inflected forms based on the relevant verb stem.
2. Present (*buys/is buying*), imperfect (*was buying*) and perfect (*has bought*) are formed as follows:
   - Present: Present tense of 'be' (*bod*) + Progressive aspect marker *yn* + Verbal noun
   - Imperfect: Imperfect tense of *bod* + *yn* + Verbal noun
   - Perfect: Present tense of *bod* + Perfect marker *wedi* + Verbal noun

Correlating this distinction with Section 1, categories 1. occur after the affirmative marker *y* in Type 1 complements, while 2. all occur with the *bod*-construction, where the first element of the complement, the verbal noun *bod*, 'incorporates' the complementizer (Section 1.3.1. 2.). It should be noted, since the complement marker can only be the verbal noun *bod*, that the main clause distinction between present tense (3SG *mae*) and imperfect (3SG *oedd*) is neutralized in such complements. A similar syntactic distinction is made with Type 2 complements, the categories in 1. taking the complementizer *a*, those in 2. the question forms of the present tense of *bod* (e.g., 3SG *ydy*). As we saw in Section 1.3.2, negation simplifies everything here, *na* being followed by the appropriate inflected tense in all cases 1. and 2.. As we shall see in Section 4, however, the scenario outlined here is a little too neat in practice principally because of two developments: *y*-complements have become largely confined to the future and conditional and marginal-

6 For an overview of tense in Welsh, see King (1993: 135–141).

ized with the past (preterite) tense; the *bod*-construction has begun to spread from affirmative to negative complements.

3 Semantics of complement clauses in Celtic

3.1 Two Types

The formal aspect of Celtic complementizers has already introduced us to the idea of there being two basic types of complement clauses based on the semantics of the main clause predicate. Based on the evidence of Germanic languages, Nordström & Boye (this volume) argue that Type 1 is epistemically neutral in that it typically presents the complement proposition from the vantage point of the predicate as the one possibility to be entertained, typically in complements after predicates that report speech or express a positive opinion as to factual status. Type 2 on the other hand presents the complement as 'open' in respect of factual status, as containing two polar possibilities, 'yes' or 'no'. Type 2 complements typically occur after interrogative, negative proposition attitude and negative semifactive (knowing, learning) predicates. Expressing as they do epistemic uncertainty, Type 2 complementizers introduce a variety of constructions other than indirect questions; but the point that is most relevant to Celtic is that all such constructions of epistemic uncertainty are clearly encoded as indirect polar questions.

3.2 Minimal Pairs

This basic semantic distinction can be further considered by the observation of minimal pairs. These can be most clearly examined in the case of the semifactive predicates ('know', 'find out'), utterance predicates ('say', 'tell'), as well as desiderative predicates ('would like'). Relevant parameters here are contrasts related to the sharing of information (Section 3.2.1), futurity and modalization (Section 3.2.2) and volitionality (Section 3.2.3). Sections 3.2.4 and 3.2.5 treat of a couple of special cases in Irish.

3.2.1 Sharing of information

For Welsh, take a basic sentence (12), comparing it with (13) and (14):

(12) *Mae* *hi* *fan hyn*
be.PRS.3SG 3SG.F here
'She is here'

(13) *Dw* *i* *'n* *gwybod* [***bo*** *hi* *fan hyn*]
be.PRS.1SG 1SG PROG know.VN **COMP** 3SG.F here
'I know that she is here'

(14) *Dw* *i* *'n* *gwybod* [***ydy*** *hi* *fan hyn*]
be.PRS.1SG 1SG PROG know.VN **COMP.be.PRS.3SG** 3SG.F here
'I know if she is here' (but other participants don't necessarily)

Here the same predicate form ('I know') can take both a Type 1 and a Type 2 complement. With the positive predicate of knowing in the present tense, Type 1 is the default setting: unless there is evidence to the contrary, the audience accepts that she is/was there. As the gloss to (14) suggests, the use of a Type 2 complement indicates that the information in the predicate, although known to the subject, is not universally shared by all participants in the discourse context. Consistent with the description above, (14) offers two possibilities (either she is/was there or not).

Syntactically, the two examples illustrate an important point made for Welsh in the previous section: when the complement verb is a present or imperfect form of the verb 'to be', the normal inflected finite form *mae* (see [12]) cannot be used but must be replaced by the verbal noun form *bod*, here in its more colloquial complementizer form *bo*. As mentioned above, these clauses are treated as *syntactically finite* and (13) illustrates one reason why: the subject of *bo* is the usual subject pronoun 3SG.F *hi* 'she' and it occurs here in its normal finite position after the verb, as with *mae hi* in (12) and *ydy hi* in (14) above. This could not occur in a true non-finite clause.[7]

With (13) above, compare the following (15) and (16) featuring the inflected future tense:

(15) *Dw* *i* *'n* *gwy-bod* [[***y***] **bydd** *hi* *fan hyn*]
be.PRS 1SG PROG know-VN (**COMP**) **be.FUT.3SG** 3SG.F here
'I know that she will be here'

7 For an outline of the various reasons for regarding such constructions as syntactically finite, see Borsley et al. (2007: 78–81). I will be returning to this issue *passim*.

(16) *Dw* *i* *'n* *gwy-bod* [[***a***] **fydd** *hi* *fan hyn*]
be.PRS 1SG PROG know-VN (**COMP.Q**) **be.FUT.3SG** 3SG.F here
'I know if she will be here'

The same semantic-pragmatic principles underlie this contrast as with (13) and (14) previously, but its syntactic realization is clearly different. In (15) because the tense is future, the usual inflected but unmutated form (*bydd*) must be used after Type 1 *y* 'that'; in (16) in contrast lenition (the soft mutation) is mandatory after the interrogative *a* (*fydd*). The use of parentheses in the two examples is to indicate that in both cases the complementizer is usually deleted in speech. Thus the only syntactic contrast present is generally the presence or absence of the mutation.

From such examples we would say that the question marker *a* fulfills the same function as English *if* in Welsh complementation: it cancels the normally positive implications of a predicate (Noonan 1985: 104–105).

3.2.2 Futurity and modalization

Consider now for Irish the semi-factive *faigh amach* 'find out' in the future tense form (irregular *gheobhaidh amach*):

(17) *Gheobhaidh* *sí* *amach* [***gur*** *imigh* *siad*]
find.FUT 3SG.F out **COMP.PST** leave.PST 3PL
'She will find out that they left'

(18) *Gheobhaidh* *sí* *amach* [***ar*** *imigh* *siad*]
find.FUT 3SG.F out **COMP.Q.PST** leave.PST 3PL
'She will find out if/whether they left'

Even though futurity inherently lacks the certainty of present or past, there is still the pragmatic question of relative confidence in the proposition. Type 1 (17) entertains only one possibility for the complement in terms of the speaker's expectation (they have left). Type 2 (18), on the other hand, leaves the matter open. The same can apply when the predicate is modalized, as with the following Irish pair (note that Irish does not distinguish formally between 'should' and 'should have'):

(19) *Ba* *cheart* *go* *mbeadh* *a* *fhios* *agam* [***go*** *raibh* *sé*
be.COND right COMP be.COND its knowing at.1SG **COMP** be.PST.DEP 3SG.M
ann]
here
'I should have known that he was here'

(20) *Ba* *cheart* *go* *mbeadh* *a* *fhios* *agam* [***an*** *raibh* *sé*
be.COND right COMP be.COND its knowing at.1SG **Q** be. PST.DEP 3SG.M
ann]
here
'I should have known if/whether he was here'

Once again, the distribution of information is clearly relevant here: (19) accepts that he *was* here; (20) does not but leaves it open.

We find the same situation with predicates of saying and telling. As we will see in the next section, verbs of saying especially when they report speech in the past, take Type 1 complements. Again, however, futurity in the predicate can induce a Type 1/Type 2 contrast ([21] and [22] for Welsh):[8]

(21) *Paid* *deud* *wrtha* *i* [***bo*** *nhw* *fan hyn*]
NEG.IMP say.VN to.1SG 1SG **COMP** 3PL here
'Don't tell me that they are here' (I'd rather they weren't)

(22) *Paid* *deud* *wrtha* *i* [***ydyn*** *nhw* *fan hyn*]
NEG.IMP say.VN to. 1SG 1SG **COMP.Q.be.3PL** 3PL here
'Don't tell me if they are here' (I'd rather not know)

The glosses elucidate the pragmatic distinction here. Type 2 is again the more "open", in (22) they might or might not be here; with (21) I suspect (confirming my worst fears) that they are.

3.2.3 Volitionality

For Irish, a somewhat different situation arises with the predicate adjective *maith* 'good' which is used in a modalized nominal construction (with the copula *ba* 'would be', see Section 2.4) as a desiderative predicate taking a Type 1 complement as in (23). Alongside this, however, we also have (24):

8 For the negated imperative, Welsh uses the defective verb *peidio* 'stop' which has solely negating content and appears only as an imperative (*paid*, plural *peidiwch*) and a verbal noun (*peidio*). The imperative forms are followed by the verbal noun of the main verb (King 1993: 227; Borsley et al. 2007: 270).

(23) *Ba* *mhaith* *leis* [***go*** *mbeadh* *a* *bhean* *ann*]
be.COND good with.3SG.M **COMP** be.COND POSS.3SG.M wife here
'He would like his wife to be there'
(Ó Siadhail 1989: 260)

(24) *Ba* *mhaith* *leis* [***dá*** *mbeadh* *a* *bhean* *ann*]
be.COND good with.3SG.M **CONJ.if** be.COND POSS.3SG.M wife here
'He would like (it) if his wife were there'

In (24), *go* is replaced by the conditional protasis marker *dá* 'if'. Clearly, neither example carries a factive implication for the complement, as is clear also from the conditional mood in the respective subordinate clauses. The difference between (23) and (24) is the degree of control exercised by the subject of the main clause on the subordinate clause. That this control is higher in the case of (23) can be shown syntactically. An alternative to (23), but not to (24), is to use a non-finite complement with the verbal noun *bheith* (25):

(25) *Ba* *mhaith* *leis* [*a* *bhean* *a* *bheith* *ann*]
be.COND good with.3SG.M POSS.3SG.M wife PRT.to be.VN here
'He would like his wife to be there'
(Ó Siadhail 1989: 260)

In addition, the order of clauses can be reversed in (24) but not in (23), although a resumptive pronoun *é* 'it' is now added:

(26) [*Dá* *mbeadh* *a* *bhean* *ann*] *ba* *mhaith* *leis* *é*
CONJ.if be.COND POSS.3SG.M wife there be.COND good with.3SG.M 3SG.M
'If his wife were there, he would like it'

Thus the *go*-clause is more semantically and syntactically integrated with its governing clause (it must be extraposed), while the *dá*-clause syntactically behaves more like an adjunct clause, in a freer relationship with the main clause, a condition clause in effect. Nonetheless, the *dá*-clause passes a syntactic test for a complement, as it counts as a subject argument of the main predicate, it fills the same argument slot.

3.2.4 The Irish predicate *amhras*

The Irish predicate *amhras* has a range of meanings from 'doubt', to 'suspicion', 'guess' and speculation', as in (27) for example. Accordingly, we find the minimal pair represented by (28) and (29):

(27) *Tá* *amhras* *orm* *faoi*
be.PRS doubt on.1SG about.3SG.M
'I'm doubtful about it'

(28) *Tá* *mé* *in* *amhras* [***go*** *bhfuil* *an* *ceart* *agat*]
be.PRS 1SG PREP.in doubt **COMP** be.PRS.DEP ART right at.2SG
'I suspect that you're right'

(29) T*á* *mé* *in* *amhras* [***an*** *bhfuil* *an* *ceart* *agat*]
be.PRS 1SG PREP.in doubt **COMP.Q** be.PRS.DEP ART right at.2SG
'I am doubtful whether you are right'
(Ó Dónaill 1977 under *amhras*, 41)

Here the choice of complementizer seems to elucidate the particular nuance of the predicate, which involves a continuum of meaning from doubt to something less than a positive propositional attitude (think, believe). Historically *amhras* continues Old Irish *amaires* 'lack of faith or belief': Type 1 *go* veers towards the positive end (i.e. 'sort of think') of the spectrum while Type 2 is at the negative end ('doubt', almost 'wonder if').[9] So we might paraphrase respectively:

(28) 'I sort of think that you're right' (i.e. Type 1)

(29) 'I sort of think (:) are you right (?)' (i.e. Type 2: I wonder if you're right)

3.2.5 Irish *má*

As is now apparent, in both Irish and Welsh the equivalent of the protasis markers used as a Type 2 complementizer in other European languages is the polar question marker. We have seen above that the Irish protasis marker *dá* 'if' (used in hypothetical conditionals) can also be used in a complement-like structure although the clause shows characteristics of an adjunct. Irish has a second protasis marker *má* 'if' (used in past, present and future conditionals) which again plays a marginal role in complementation.

In certain fixed or stereotyped phrases, *má* contrasts with a Type 1 complementizer. An example is the fixed phrase *beag*/*ba bheag* 'it is/was (by) little', which is the equivalent of English *barely*, *scarcely*, and followed by a finite complement. We have the following minimal pair: (30) comments on an underlying

9 See de Bhaldraithe (1956/7) for discussion of a western Irish dialect here.

factual proposition 'they didn't win' while (31) comments on an underlying 'they won':

(30) *Ba* *bheag* [***nár*** *bhuaigh* *siad*]
be.PST little **COMP.NEG.PST** PST\win 3PL
(It was (by) little that they did not win)
'They almost won'

(31) *Ba* *bheag* [***má*** *bhuaigh* *siad*]
be.PST little **CONJ.if** PST\win 3PL
(It is (by) little if they won)
'They barely won'

If the premise underlying (31) is accepted as factual, why would the conditional marker *má* be used? Its purpose here appears to be attenuation of the complement proposition, underscoring the fact that if it *is* so, it is *barely* so. The speaker, while accepting the complement as factual, distances herself from it or attenuates her commitment to it. A second instance with *má* occurs after the exclamative *damnú* 'damnation':

(32) *Damnú* [***go*** *bhfuil* *a* *fhios* *agam*]
Damnation **COMP** be.PRS.DEP POSS.3SG.M knowing at.1SG
'I damned well know!' (Damnation that I know).

(33) *Damnú* [***má*** *tá* *(a)* *fhios* *agam*]
Damnation **COND.if** be. PRS (POSS.3SG.M) knowing at.1SG
'I am damned if I know'
(Ó Siadhail 1989: 326)

The implication of (32) is positive: 'I know!' and I am going to emphasize the fact that I do. Therefore it is no surprise to find Type 1 *go* used here. In the case of (33) the implication is the reverse, 'I most certainly don't know' (and if I *did* know, the devil could take my soul, compare German *hol mich der Teufel wenn*). The effect of *má* here is to attenuate the positive implications of the complement itself: it is an expressive way, in effect, of negating it.

The minimal pairs which contrast Type 1 and Type 2 show the validity of the proposed description. In complements introduced by Type 1 complementizers, one possibility is presented for the complement; this is generally assumed to be factual but in non-factual modalized contexts can otherwise reflect the predicate's greater desire to control the complement, which is therefore syntactically more integrated (Givón 1995: 126). Type 2 complementizers on the other hand, imply that there is more than one possibility for the status of the complement. The use of a Type 1 complementizer with an indicative verb presupposes the factual status of the complement; the use of Type 2 does not, but neither does it

exclude its possible factuality. The Irish protasis marker *má* can also be used to attenuate the force of a complement proposition otherwise regarded as factual. This is germane to Type 2 complementizer usage whereby factuality is in some way attenuated but not excluded.[10]

However, it must be noted that *go* does not, in and of itself, indicate that the complement is factual: this is a function of the tense or mood of the complement clause as shown above by the conditional mood. Consider once again (23) above, 'He would like his wife to be here' where Irish *go* is followed by an irrealis complement verb form. The point is that any finite verb form permitted in an Irish subordinate clause can be preceded by *go*, a fact highlighting its semantic neutrality. This also applies to Welsh *y* 'that', except that, as we have seen, it has to be followed by an inflected finite form. But germane to the point here is that this can be an irrealis conditional (Note: Welsh adjectives when used predicatively are preceded by a predicate marker *(y)n*):

(34) *Mae* *'n* *sicr* [***y*** byddai *hynny* *'n* *beryglus*]
be.PRS.3SG PRED certain **COMP** be.COND.3SG this PRED dangerous
'It's certain that this would be dangerous'

3.3 Combinability of complementizers

There are also facts relating to the behavior of the complementizers themselves that support the above analysis as types 1 and 2.

First of all, the two types cannot combine. On the other hand, certain combinations arise within Type 2. Irish has a wh-question marker *cé acu* ('which of them') which governs a following relative clause (35) in the standard language (GGBC 1960: 349–350). However in more colloquial usage it can be combined with the polar marker *an* as in (37). Compare the following variants where (35) simply uses the polar question marker, whereas with *cé acu* the two polarities have to be spelt out:

(35) *Níl* *a* *fhios* *agam* [***an*** *ól-ann* *sí*]
NEG.be.PRS its knowing at.1SG **COMP.Q** drink-PRS 3SG.F
'I don't know if she drinks'

10 The term non-exclusion of factuality was used in relation to similar complements by Benjamin Fagard, Julie Glikman and Paola Pietrandrea in their presentation 'Syntactic and semantic aspects of Romance complementizers' given at the LANCHART workshop on Semantic Functions of Complementizers in European Languages, University of Copenhagen, October 29th 2011. See further this volume.

(36) *Níl* *a* *fhios* *agam* [***cé acu*** ***a*** *ól-ann* *sí*] *nó* *nach*
NEG.be.PRS its knowing at.1SG **COMP.Q** **REL** drink-PRS 3SG.F or NEG
n-ól-ann
drink-PRS
'I don't know if she drinks or not'

(37) *Níl* *a* *fhios* *agam* [***cé acu*** ***an*** *ól-ann* *sí*] *nó*
NEG.be.PRS its knowing at.1SG **COMP.Q** **Q** drink-PRS 3SG.F or
nach *nól-ann*
NEG drink-PRS
'I don't know if she drinks or not'

There is no difference in meaning between (36) and (37), the version with the relativizer *a* is standard written Irish while that with the polar question marker *an* is the more colloquial of the two, combining harmonically the two Type 2 question markers which otherwise appear separately as in (35) and (36). Consider further (38), where in dialectal usage, the negative interrogative can be replaced by the negative protasis marker *mura* (*mara*) 'if not' (see Ó Siadhail 1989: 321):

(38) *Níl* *a* *fhios* *agam* [***an*** *ól-ann* *sí* *nó*
NEG.be.PRS its knowing at.1SG **COMP.Q** drink-PRS 3SG.F or
mara *nól-ann*]
NEG.CONJ.if drink-PRS
'I don't know if she drinks'

I noted above the marginal position of protasis markers in Irish complementation: the above example, however, shows a conceptual coherence between interrogative and conditional markers since both involve the 'non-exclusion of factuality' (see note 11), both leave open the two polar possibilities for the complement. Consider further (39) for Welsh:

(39) *Dw* *i* *ddim* *yn* *sicr* [***ydy*** *hi* *'n* *siarad*
be.PRS.1SG 1SG NEG PRED sure **COMP.Q.be.PRS.3SG** 3SG.F PROG speak.VN
Cymraeg] (*neu* *beidio*)
Welsh (or not)
'I'm not sure if she speaks Welsh or not'
(King 1993: 305)

Some speakers introduce the protasis marker *os* 'if' into the construction before the polar question marker as in (40):

(40) *Dw* *i* *ddim* *yn* *sicr* [***os*** ydy *hi* *'n*
be.PRS.1SG 1SG NEG PRED sure **CONJ.if** Q.be.PRS.3SG 3SG.F PROG
siarad *Cymraeg*]
speak.VN Welsh
'I'm not sure if she speaks Welsh'

According to King, this is regarded as substandard' by many speakers and is assumed to represent English influence; according to Koch (p.c), the *os* construction is "very common" in speech but is frowned upon by grammarians.[11] Nonetheless the combination of *os* and *ydy* here is semantically harmonic as both signal a level of epistemic uncertainty.

4 Distribution

4.1 Type 1

4.1.1 Positive Propositional Attitude

This is the most numerous class of complement-taking predicates in both languages and consists of the types 'think', 'believe', 'be certain/necessary/probable and possible'.

4.1.1.1 Irish

Irish has two verbs 'to be', a substantive verb *tá* which typically links the subject to a predicate adjective or adverbial, and a copula *is* linking the subject with a predicate noun (see Section 2.4). Let us first consider *tá* in the basic sentences, affirmative (41) and negative (42):

(41) *Tá* *cathaoir* *anseo*
be.PRS chair here
'There is a chair here'

11 Koch (p.c.) also points out to me that there may be an issue of register here which is complicated by speakers' knowledge of English: in less formal registers of the latter *if* is generally favoured in such constructions over *whether*; the popularity of *os* over *a* may reflect a similar situation in Welsh under English influence.

(42) *Níl* *cathaoir* *anseo*
NEG.be.PRS chair here

The verb *tá* 'is' is one of a handful in Irish requiring a special dependent form after certain particles, including Type 1 complementizers *go* (43) and *nach* (44). In both cases, the dependent form *bhfuil* must be used after the complementizer. Both (43) and (44) show the positive proposition predicate *síl* 'think' (Ó Siadhail 1989: 253):

(43) *Síl-im* [***go*** *bhfuil* *cathaoir* *anseo*]
think-1SG **COMP** be.PRS.DEP chair here
'I think there is a chair here'

(44) *Síl-im* [***nach*** *bhfuil* *cathaoir* *anseo*]
think-1SG **NEG.COMP** be.PRS.DEP chair here
'I think there isn't a chair here'

As discussed briefly in Section 2.4 above, Irish has a second verb 'be' (copula) which links the subject with a predicate noun:

(45) *Is* *múinteoir-í* *iad*
be.PRS teacher-PL 3PL
'They are teachers'

(46) *Ní* *múinteoir-í* *iad*
PRS.be.PRS teacher-PL 3PL
'They are not teachers'

In the case of the copula *is*, the complementizer *go* incorporates the verbal form to give a present form *gur* (47); the form of the copula with negative complementizer is *nach* (48):

(47) *Is* *cinnte* [***gur*** *múinteoir-í* *iad*]
be.PRS certain **COMP** teachers-PL 3PL
'It is certain that they are teachers'

(48) *Is* *cinnte* [***nach*** *múinteoir-í* *iad*]
be.PRS certain **NEG.COMP** teachers-PL 3PL
'It is certain that they are not teachers'

Included here are various expressions of epistemic modality which almost always take a finite complement governed by *go* or *nach*:

(49) *Caith-fidh sé* [***go*** *bhfuair tú coláiste*]
must-FUT 3SG.M **COMP** get.PST 2SG college
(=It must be that you got college)
'You must be college educated'
(Ó Siadhail 1989: 272)

(50) *Thioc-fadh leis* [***go*** *bhfuil sé fíor*]
come-COND to.3SG.M **COMP** be.PRS.DEP 3SG.M true
(=It would come to it that it is true)
'It could be true'
(Ó Siadhail 1989: 273)

(51) *Ní féidir* [***go*** *bhfuil Bríd ag foghlaim Fraincis*]
NEG.COP possible **COMP** be.PRS.DEP Bríd PROG learning French
'It can't be the case that Bríd is learning French'
(Ó Siadhail 1989: 290)

(52) *B'fhéidir* [***nach*** *mbeadh sé ann*]
NEG.be.COND.possible **NEG.COMP** be.COND 3SG.M there
'He might not be there'

4.1.1.2 Welsh

According to the syntactic principles outlined in sections 1 and 2 above, the complementizer *y* is followed by an inflected future or conditional. In practice *y* is generally omitted in speech, as in (55):

(53) *Swn i 'n meddwl* [***y*** *dôn nhw wedyn*]
be.COND.1SG 1SG PROG think.VN **COMP** come.FUT.3PL 3PL later
'I should think they'll be along later'
(King 1993: 306)

(54) *Efallai* [***y*** *bydda i 'n ffonio wedyn*]
perhaps **COMP** be.FUT.1SG 1SG PROG phone later
'Perhaps I'll call later/I might call later'
(King 1993: 261)

(55) *Mae hi 'n debyg bydd glaw trwm yn yr haf*
be.PRS.3SG 3SG.F PRED likely be.FUT.3SG rain heavy PREP.in ART summer
'It is likely (that) there will be heavy rain in the summer'
(Borsley et al. 2007: 101)

The negative complementizer *na* is followed by any inflected verb form in the complement. Note the lenition of *dylech* to *ddylech*:

(56) *Dw* *i* *'n* *siwr* [***na*** *ddylech* *chi* *ddeud* *wrtho*
be.PRS.1SG 1SG PRED sure **NEG.COMP** should.COND 2PL say.VN to.3SG.M
fe]
3SG.M
'I'm sure you shouldn't tell him'
(King 1993: 306)

In focus constructions, the complementizer is *mai* (57), while the negative form is *nad* (58):

(57) *Mae* *'n* *amlwg* [***mai*** *chi* *sy* *ar fai*]
be.PRS.3SG PRED obvious **COMP** 2PL REL.PRS to blame
'It is clear that **you** are to blame' (It is clear that it is **you** who are to blame').
(King 1993: 307)

(58) *Dw* *i* *'n* *siwr* [***nad*** *ni* *sy* *'n* *gyfrifol*]
be.PRS.1SG 1SG PRED sure **NEG.COMP** 1PL REL.PRS PRED responsible
'I'm sure that **we** are not responsible' (I am sure that it is not we who are responsible).
(King 1993: 309)

If the verb of an affirmative complement is a present or imperfect tense form of the verb *bod* 'be', then the verbal noun *bod* replaces the inflected form and incorporates the marking of the complement. In (59) *bod* appears lenited as *fod*:

(59) *Rhaid* [***fod*** *Elen* *yn* *y* *dre*]
Need **COMP**.be.VN Elen PREP.in ART town
'Elen must be in the town'
(King 1993: 212)

(60) *Mae* *Aled* *yn* *credu* [***bod*** *Elen* *yn* *darllen* *y* *llyfr*]
be.PRS.3SG Aled PROG believe.VN **COMP**.be.VN Elen PROG read.VN ART book
'Aled believes that Elen is/was reading the book'
(Borsley et al. 2007: 77)

Compare the initial verb form of the two clauses in (60) where in effect *bod* replaces an original inflected *mae* as in (61):

(61) ***Mae*** *Elen* *yn* *darllen* *y* *llyfr*
be.PRS.3SG Elen PROG read.VN ART book
'Elen is/was reading the book'

The somewhat neat division (Section 2.5. above) for purposes of complementation between simple inflected forms and periphrastic ones is skewed by developments with the preterite (past). According to the schema of Section 2.5 we would expect as in (62) a *y*-complement (the actual sentence given in Borsley et al. [2007] is [64] below, [62] and [63] are the possible variants):

(62) *Mae* *Aled* *yn* *sicr* [***y*** *aeth* *Mair*]
be.PRS.3SG Aled PRED sure **COMP** go.PST.3SG Mair
Aled is certain that Mair went/has gone'

According to Borsley et al. (2007: 82), however, while some speakers do permit the morphological past tense form in this environment, either the *bod*-construction with the periphrastic perfect (63) or the construction using the preposition *i* 'to' (64) are preferred (see sections 1.3.1 and 2.5):

(63) *Mae* *Aled* *yn* *sicr* [***bod*** *Mair* *wedi* *mynd*]
be.PRS.3SG Aled PRED sure **COMP.be.VN** Mair PFV go.VN
'Aled is certain that Mair went/has gone'

(64) *Mae* *Aled* *yn* *sicr* [***i*** *Mair* *fynd*]
be.PRS.3SG Aled PRED sure **COMP.PREP.to** Mair go.VN
'Aled is certain that Mair has gone'
(Borsley et al. 2007: 83)

Example (63) substitutes the periphrastic perfect for the inflected past and hence we have the expected *bod*-construction (*wedi* 'after' functioning here as the perfect marker). In (64), the complement verb form is also expressed by the verbal noun (here lenited) while its semantic subject is governed by the preposition *i* 'to' as indirect object. This is a construction that can only be used if the complement proposition is anterior to that of the main clause. Examples (63) and (64) are therefore semantically equivalent, but according to King (1993: 307), the *bod* + *wedi* construction is the more common; according to Koch (p.c), the two are "close to equivalent", the past perfect being better characterized in (63).

4.1.2 Predicates of knowledge and acquisition of knowledge

The most common complement-taking predicates here are 'know', 'find out' and 'forget'.

4.1.2.1 Irish

Irish has no verb 'know' but uses a nominal phrase as presented in (65), 'its knowing is at me' = 'I have its knowing':

(65) *Tá* *a* *fhios* *agam* [***nach*** *bhfuil* *sé* *ann*]
be.PRS POSS.3SG.M knowing at.1SG **NEG.COMP** be.PRS.DEP 3SG.M there
'I know that he is not there'

In the verb *tá* 'be', the past tense is *bhí* but it has a special dependent form *raibh* which is used after conjunctions and complementizers such as *go*:

(66) *Fuair mé amach* [***go*** *raibh sí sásta*]
find.PST 1SG out **COMP** be.PST.DEP 3SG.F happy
'I found out that she was happy'

As in the case of 'know' above, there is no verb 'forget', rather 'make a forgetting' (or a 'mistake', as the word *dearmad* may also be translated):

(67) *Rinne muid dearmad* [***gur*** *imigh siad*]
make.PST 1PL forgetting **COMP.PST** leave.PST 3PL
'We forgot that they had left'

Note here that the past form of the complementizer in (67), *go* becomes *gur* before regular past tense forms.

4.1.2.2 Welsh

The formal and syntactic principles have been outlined in sections 1.3.1 and 2.5 and illustrated in Section 4.1.1.2 above. A *y*-complement takes an inflected form and as we know the complementizer can be omitted (69): (68) is future while (69) contains the marginal (in complementation) inflected past:

(68) *Dw i 'n gwybod* [***y*** *bydd-an nhw 'n hwyr*]
be.PRS.1SG 1SG PROG know.VN **COMP** be.FUT-3PL 3PL PRED late
'I know that they will be late'

(69) *Dw i 'n gwybod* [*gwel-odd Siôn y gêm*]
be.PRS.1SG 1SG PROG know.VN see.PST-3SG Siôn ART game
'I know that Siôn saw the game'
(Borsley et al. 2007: 82)

As Borsley et al. say of the use of the inflected past in complements (69): "some speakers do permit this, but in a number of spoken varieties, as well as in literary Welsh, the past tense is ungrammatical" (Borsley et al. 2007: 82). (70) illustrates the focus construction marked by the complementizer *mai*:

(70) *Dw i 'n gwybod* [***mai*** *nhw sy 'n hwyr*]
be.PRS.1SG 1SG PROG know.VN **COMP** 3PL REL.PRS late
(I know that it is they who are late)
'I know that **they're** late'

As discussed in Section 4.1.1.2, the constructions in (71) and (72) are now both preferred to (69) above when the complement is anterior to the main predicate. Note also that *'r* (*yr*) is the form of the definite article after a vowel as in (71):

(71) *Dw* *i* *'n* *gwybod* [***i*** *'r* *trên* *fynd*]
be.PRS.1SG 1SG PROG know.VN **COMP.PREP.to** ART train VN.go
'I know that the train went'
(King 1993: 307)

(72) *Dw* *i* *'n* *gwybod* [***fod*** *y* *trên* *wedi* *mynd*]
be.PRS.1SG 1SG PROG know.VN **COMP.VN.be** ART train PFV go.VN
'I know that the train went'
(King 1993: 307)

Otherwise, the *bod*-construction appears as expected when the complement verb is affirmative and is a present or imperfect form of *bod* 'to be' (73):

(73) *Dw* *i* *'n* *gwybod* [***bod*** *nhw* *'n* *hwyr*]
be.PRS.1SG 1SG PROG know.VN **COMP.VN.be** 3PL PRED late
'I know that they are late'
(King 1993: 309)

When the complement verb is negative, the appropriate tensed form is used. In (74), the form *ddim* intensifies the negation and may be omitted:

(74) *Dw* *i* *'n* *gwybod* [***nad*** *ydy* *Ffred* *(ddim)*
be.PRS.1SG 1SG PROG know.VN **NEG.COMP** be.PRS.3SG Fred (NEG)
yn *dod*]
PROG come.VN
'I know that Fred is not coming'
(King 1993: 309)

Here arises a further 'complication' to the neat schema presented in Section 2.5: increasingly, the *bod*-construction, negated by *ddim* as in (75), is making inroads as an alternative to (74) above:

(75) *Dw* *i* *'n* *gwybod* [***fod*** *Ffred* ***ddim*** *yn dod*]
be.PRS.1SG 1SG PROG know.VN **COMP.be.VN** Fred **NEG** PROG come.VN
'I know that Fred isn't coming'
(King 1993: 309)

Note that in (75), unlike in (74), *ddim* carries the negation and may not be omitted.

4.1.3 Utterance predicates

Included here are the verbs 'say', 'tell' and 'report'.

4.1.3.1 Irish

The main clause predicate in (76) and (77) is the irregular verb *abair* 'say' (present stem *deir*, past stem *dúirt*), while (77) again shows the past tense dependent form *raibh* after the negative complementizer *nach;* (78) is an example of the main predicate *inis* 'tell':

(76) *Deir-im* [***gur*** *ceart Colm*]
say.PRS-1SG **COMP.be.PRS** right Colm
'I say that Colm is right'
(GGBC 1960: 350)

(77) *Dúirt sé liom* [***nach*** *raibh an scéal fíor*]
tell.PST 3SG.M to.1SG **NEG.COMP** be.PST.DEP ART story true
'He told me that the story was not true'
(based on Ó Siadhail 1989: 270)

(78) *D'inis sé dhom* [***go*** *raibh an doras briste*]
AFF.tell.PST 3SG.M to.1SG **COMP** be.PST.DEP ART door broken
'He told me that the door was broken'
(Ó Siadhail 1989: 275)

4.1.3.2 Welsh

Apart from positive propositional attitude predicates, predicates of saying are the most prominent in taking the various Type 1 complement structures outlined above, here exemplified by the main clause predicate *dweud*/*deud* 'say' ([79] to [83]):

(79) *Wed-es i* [***y*** *dylai fo fod wedi codi*]
say.PST-1SG 1SG **COMP** should.COND.3SG 3SG.M be.VN after rise.VN
'I said he ought to have got up'
(King 1993: 231)

(80) *Wed-odd hi* [***na*** *gâi hi byth swydd eto*]
say.PST-3SG 3SG.F **NEG.COMP** get.COND.3SG 3SG.F ever job other
'She said she would never get another job'
(King 1993: 195)

(81) *Mae e 'n dweud* [***mai*** *Gerwyn naeth e*]
be.PRS.3SG 3SG.M say.VN **COMP** Gerwyn do.PST.3SG it
(he says that it is Gerwyn who did it)
He says that **Gerwyn** did it'
(based on King 1993: 310)

(82) *Mae* *e* *'n* *deud* [***bod*** ***n****hw* *fan hyn*]
be.PRS.3SG 3SG.M PROG say.VN **COMP.be.VN** 3PL here
'He says that they are here'
(King 1993: 304)

(83) *Wed-odd* *nhw* [***i*** *Mair* *fynd* *adre*]
say.PST-3SG 3PL **COMP.PREP.to** Mair go.VN home
'They said that Mair went home'

Once again if the complement verb in (83) is negated, then the inflected past form (here *aeth*) is preferred (84); *na* becomes *nad* before a following vowel:

(84) *Wed-odd* *nhw* [***nad*** *aeth* *Mair* *adre*]
say.PST-3SG 3PL **NEG.COMP** go.PST.3SG Mair home
'They said that Mair did not go home'

It was stated above (Section 1.3.1) that there are reasons for regarding constructions like (82) and (83) as syntactically finite. Compare (83) with (84) for example: clearly (83) though formally non-finite, slots into what is essentially a finite syntactic paradigm as it must take a tensed verb form when negative. Similarly, if the complement verb were changed to future or conditional then the verb form would have to be finite, as in (79) and (80). Another piece of evidence are the facts relating to pronominal arguments. Compare below literary (85) and colloquial Welsh ([86] and [87]). Note the various options provided for in the English translation:

(85) *Dywed-odd* [*ei* ***fod*** *yn* *dod*]
say.PST-3SG POSS.3SG.M **COMP.be.VN** PROG come.VN
'He said that he comes/is coming/was coming/used to come'
(Ball & Fife 1993: 342)

(86) *Wed-odd* *e* [***fod*** *e* *'n* *dod*]
say.PST-3SG 3SG.M **COMP.be.VN** 3SG.M PROG come.VN
'He said that he comes/is coming/was coming/used to come'

(87) *Wed-odd* *e* [***bo*** *fe* *'n* *dod*]
say.PST-3SG 3SG.M **COMP** 3SG.M PROG come.VN
'He said that he comes/is coming/was coming/used to come'

In the literary version the verbal noun takes the pronominal argument appropriate to its nominal status: a possessive ('his coming'). In the colloquial version (86), however, the verbal noun *bod* takes the same subject marking (*e*) as the matrix finite verb *wedodd* but residually shows the soft mutation which occurs after the 3SG.M possessive (as in [85]). In the most colloquial version (87), the form

of *bod* is *bo* which is not subject to any initial mutation and is again followed by the appropriate finite subject marking (here *fe* after a vowel).[12]

4.1.4 Perception predicates

Included here are the verbs 'see', 'hear', and 'notice'.

4.1.4.1 Irish

Comparison of examples (88) and (89) below illustrates the mandatory extraposition of finite complements. In (88) the object noun phrase (*an bhean*) can come directly after the sequence verb-subject even though it breaks up the verb phrase (*thug sé faoi deara* 'he noticed'); in (89) however the noun *clause* cannot (Ó Siadhail 1989: 275):

(88) *Thug* *sé* [*an* *bhean*] *faoi* *deara*
take.PST 3SG.M ART woman under notice
'He noticed the woman'

(89) *Thug* *sé* *faoi* *deara* [***go*** *raibh* *an* *bhean* *ann*]
take.PST 3SG.M under notice **COMP** be.PST.DEP ART woman there
'He noticed that the woman was there'

The same principle can be exemplified by other perception predicates which can be followed both by a verbal noun expressing a state-of-affairs ([90]/[92]) and also by a full-fledged complement expressing a proposition ([91]/[93]). Compare the following with the verbs 'hear' (*clois*, irregular past *chuala*) and 'see' (*feic*, irregular past *chonaic*):

(90) *Chonaic* *mé* [*ag* *damhsa*] *é*
see.PST 1SG PROG dance.VN 3SG.M
'I saw him dancing'

(91) *Chonaic* *mé* [***go*** *raibh* *sé* *ag* *damhsa*]
see.PST 1SG **COMP** be.PST.DEP 3SG.M PROG dance.VN
'I saw that he was dancing'

12 For a succinct discussion of the reasons for regarding both verbal-noun constructions as syntactically finite, see Borsley et al. 2007: 78–81 for the *bod*-construction and 82–85 for the *i*-construction.

(92) *Chuala* *mé* *iad* [*ag* *caint*] *faoi*
hear.PST 1SG 3PL PROG talk.VN about.3SG.M
'I heard them talking about him'

(93) *Chuala* *mé* [***go*** *raibh* *siad* *ag* *caint* *faoi*]
hear.PST 1SG **COMP** be.PST.DEP 3PL PROG talk.VN about.3SG.M
'I heard that they were talking about him'

The difference between ([90]/[92]) and ([91]/[93]) is that the event in the former is *necessarily* simultaneous with the act of perception in the predicate whereas that in the latter is not (Noonan 1985: 130).

4.1.4.2 Welsh

Again there is the contrast between (94) and (95): an affirmative complement takes the *bod*-construction, while a negative complement requires an inflected finite form. The construction in (94) reflects literary influence in that the form *mod* is the nasalized product of *bod* after the elided 1sg. possessive *fy*: literally this reads something like, 'don't you notice my wearing (of) the anorak?':

(94) *Wyt* *ti* *ddim* *yn* *sylwi* [***'mod*** i *'n* *gwisgo*
Q.PRS.2SG 2SG NEG PROG notice.VN **COMP.be.VN** 1SG PROG wear.VN
anorac?]
anorak
'Don't you notice that I'm wearing an anorak?'
(Rhys Jones 1977: 207)

(95) *Fe* *glyw-es* [***nad*** *oedd* *Mr. Owen* *ddim* *yno*]
AFF hear.PST-1SG **NEG.COMP** be.IPFV.3SG Mr. Owen NEG there
'I heard (that) Mr Owen wasn't there'
(Rhys Jones 1977: 203)

4.1.5 Commentative predicates

These are predicates commenting on complement propositions assumed to be factual. They can involve emotions or judgments. Note the characteristic Celtic form of predicates expressing emotions: 'X is sad/ afraid' = 'There is sorrow/fear on X', see (97) for Irish and (100) for Welsh.

4.1.5.1 Irish

(96) *Is* *maith* *an* *rud* *é* [***go*** *bhfuil* *sé* *ann*]
be.PRS good ART thing 3SG.M **COMP** be.PRS.DEP 3SG.M there
'It's a good thing that he is there'

(97) *Tá* *brón* *orm* [***nach*** *bhfuil* *sé* *ann*]
be.PRS sorrow on.1SG **COMP** be.PRS 3SG.M here
'I'm sorry that he is not here'

(98) *Is* *é* *an* *trua* *é* [***go*** *raibh* *sibh* *ann*]
be.PRS 3SG.M ART pity 3SG.M **COMP** be.PRS.DEP 2PL there
'It's a pity that you were there'
(Ó Siadhail 1989: 261)

4.1.5.2 Welsh

(99) *Dw* *i* *'n* *falch* [***bo*** hi *'n* *darllen* *y* *llyfr*]
be.PRS.1SG 1SG PRED glad **COMP** 3SG.F PROG read.VN ART book
'I'm glad that she is reading the book'

(100) *Mae* *arna* *i* *ofn* [***fod*** *ei* *thymheredd* *hi* *'n*
be.PRS.3SG on.1SG 1SG fear **COMP**.be.VN POSS.3SG.F temperature 3SG.F PROG
codi]
rise.VN
'I'm afraid that her temperature is rising'
(Rhys Jones 1977: 265)

4.1.6 Desiderative predicates

These are predicates of the type 'hope', 'want' and 'wish'. See also Section 3.2 above on minimal pairs.

4.1.6.1 Irish

See also 3.2.3 above, examples (23) through (26).

(101) *Tá* *súil* *aici* [***nach*** *n-imeoidh* *siad*]
be.PRS hope at.SG.F **NEG.COMP** leave.FUT 3PL
'She hopes that they won't leave'

(102) *Ba* *mhaith* *leis* [***go*** *mbeadh* *a* *bhean* *ann*]
be.COND good with.3SG.M **COMP** be.COND POSS.3SG.M wife there
'He would like his wife to be there'
(Ó Siadhail 1989: 260)

4.1.6.2 Welsh

The verbal noun *gobeithio* 'hoping' is used in the manner of English *hopefully*. Note the omission of the complementizer in (103) but also how it cannot be omitted in a focus complement (104):

(103) *Gobeithio* [*bydd* *y* *tywydd* *yn* *braf*]
hope.VN be.FUT.3SG ART weather PRED fine
'I hope that the weather will be fine'
(Rhys Jones 1977: 293)

(104) *Gobeithio* [***mai*** *tywydd* *braf* *gyda ni*]
hope.VN **COMP** weather fine with.1PL
'I hope that we'll have **fine weather**'
(Rhys Jones 1977: 293)

4.2 Type 2

In both Irish and Welsh, Type 2 complementizers are polar question markers. They are found after predicates of asking, uncertainty, doubt or not-knowing. The relationship between polar question markers and complement types is most obvious after predicates of asking, which are essentially indirect questions.

4.2.1 Irish

Compare (105) and (106) with the verb *bris* 'break' where the initial *b* is nasalized to */m/* after the question particle *an*:

(105) *Bris-eann* *sé* *go* *réidh*
break-PRS 3SG.M ADV easy
'It breaks easily'

(106) ***An*** *mbris-eann* *sé* *go* *réidh?*
Q break-PRS 3SG.M ADV easy
'Does it break easily?'

Before the vast majority of verbs in the past tense, the form of the polar question marker is *ar* and the lenition of the verb is retained:

(107) *Bhris* *sé* *go* *réidh*
PST\break 3SG.M ADV easy
'It broke easily'

(108) ***Ar*** *bhris* *sé* *go* *réidh?*
Q.PST PST\break 3SG.M ADV easy
'Did it break easily?'

4.2.1.1 Predicates of asking

After predicates of asking, such direct questions are simply retained as indirect:

(109) *D'fhiafraigh* *sí* *díom* [***an*** *mbriseann* *sé* *go* *réidh*]
PST\ask 3SG.F of.1SG **COMP.Q** break.PRS 3SG.M ADV easy
'She asked me if it breaks easily'

(110) *D'fhiafraigh* *sí* *díom* [***ar*** *bhris* *sé* *go* *réidh*]
PST\ask 3SG.F of.1SG **COMP.Q.PST** PST\break 3SG.M ADV easy
'She asked me if it broke easily'

Indirect questions as complements retain the semantic structure of direct polar questions: more than one possibility is allowed for the complement (as one can answer a direct question with either 'yes' or 'no'). Accordingly, the structure of such complements can be amplified to make this explicit by also using the negative interrogative, here *nár* in the past, thereby explicitly juxtaposing the two possibilities:

(111) *D'fhiafraigh* *sí* *díom* [***ar*** *bhris* *sé* *nó*
PST\ask 3SG.F of.1SG **COMP.Q.PST** PST\break 3SG.M or
nár *bhris*]
NEG.Q.PST PST\break
'She asked me if it broke or not'

There is a further variation on (111) above, the use of the phrase *cé acu* 'which of them?' (see Section 1.3.4. above):

(112) *D'fhiafraigh* *sí* *díom* [***cé*** ***acu*** ***ar*** *bhris*
PST\ask 3SG.F of.1SG **COMP.Q** **at.3PL** **COMP.Q.PST** PST\break
sé *nó* *nár* *bhris*]
3SG.M or NEG.Q break.PST
'She asked me if it broke or not'

Examples (111) and (112) are semantically equivalent. In more standard written Irish the relativizer *a* is used instead of the question marker *ar* in (112).

Another common predicate of asking is the expression *cuir ceist ar*, literally 'put a question on' as in (111). Note that with a few irregular verbs in the past the non-past interrogative / complementizer *an* is used:

(113) *Chuir* *mé* *ceist* *air* [***an*** *raibh* *airgead* *aige*]
PST\put 1SG question on.3SG.M **COMP.Q** be.PST.DEP money at.3SG.M
'I asked him if he had money'

4.2.1.2 Predicates of uncertainty and doubt

The same structures occur after predicates expressing uncertainty, doubt or lack of knowledge ('don't know'):

(114) *Níl* *mé* *cinnte* [***an*** *raibh* *sí* *ann*]
NEG.be.PRS 1SG sure **COMP.Q** be.PRS.DEP 3SG.F there
'I'm not sure if she was there'

(115) *Cá* *bhfios* [***an*** *dtioc-fadh* *sí*]
WH knowing **COMP.Q** come-COND 3SG.F
'Who knows if she would come'
(based on *GGBC* 349)

(116) *Tá* *mé* *in* *amhras* [***an*** *bhfuil* *an* *ceart* *agat*]
be.PRS 1SG PREP.in doubt **Q** be.PRS.DEP ART right at.2SG
'I doubt if you're right'
(Ó Dónaill 1977 under *amhras*, 41)

This type of structure also occurs after the predicate *is cuma* 'it is equal / the same (whether)'. In (117) the relativizer *a* rather than the question form is used after *cé acu*:

(117) *Is* *cuma* [***cé*** ***acu*** *a-tá* *nó* *nach* *bhfuil*]
be.PRS same **COMP.Q** **3PL** REL-be.PRS or NEG be.PRS.DEP
'No matter whether or not'
(Ó Dónaill 1977 under *cuma*, 353)

4.2.2 Welsh

As we have seen in sections 1 and 2 above, the distribution of Type 2 complement types mirrors that of Type 1, a distinction being made in non-focus complements

between simple inflected verb forms (in practice future and conditional) and periphrastic tenses based on the verb *bod*. In the former instance, the polar question marker *a* (118) serves as complementizer (119), in the latter the question forms of the present tense of *bod* are used (122), inflected for person and number, so that once again verb and complement marking are incorporated in the one form. Further examples of this type are presented by (120) and (121):

(118) (***A***) *fyddai* *hi* *'n iawn?*
Q be.COND.3SG 3SG good
'Would it be OK?'

(119) *Ewch* *i* *ofyn* ***a*** *fyddai* *hi* *'n* *iawn*
go.IMP.2PL PREP.to ask.VN **COMP.Q** be.COND.3SG 3SG PRED good
'Go and ask if / whether it would be OK'
(King 1993: 309)

(120) *Dw* *i* *ddim* *yn* *siwr* [***a*** *ddylech* *chi* *ddeud*
be.PRS.1SG 1SG NEG PRED sure **COMP.Q** should.COND.2PL 2PL say.VN
wrtho *fe*]
to.3SG.M 3SG.M
'I'm not sure if/whether you should tell him'
(King 1993: 306)

(121) *Mae* *'n* *amheus* *gyda* *fi* [***a*** *aiff* *e* *'n ôl*]
be.PRS.3SG PRED doubtful with.1SG 1SG **COMP.Q** go.FUT.3SG 3SG.M back
'I doubt if he'll go back'
(Rhys Jones 1977: 267)

Where the present tense question forms of *bod* (122) mark the complement, they are inflected for person and number. In (123), the phrase *neu beidio* 'or not' makes explicit that there are two possibilities for the complement (like Irish *cé acu nó* in Section 4.2.1 above):

(122) ***Ydy*** *hi* *'n* *siarad* *Cymraeg?*
Q.be.PRS.3SG 3SG.F PROG speak.VN Welsh?
'Does she speak Welsh?'

(123) *Dw* *i* *ddim* *yn* *sicr* [***ydy*** *hi* *'n* *siarad*
be.PRS.1SG 1SG NEG PRED sure **COMP.Q.be.PRS.3SG** 3SG.F PROG speak.VN
Cymraeg *(neu beidio)*]
Welsh (or not)
'I'm not sure if she speaks Welsh (or not)'
(King 1993: 305)

(124) *Tybed* ***ydy*** *e* *'n* *bwriadu* *dod*
wonder.VN **COMP.Q.be.PRS.3SG** 3SG PROG intend.VN come.VN
'I wonder if he's intending to come'
(King 1993: 197)

(125) *Dach* *chi* *'n* *gwybod* [***ydyn*** *nhw* *'n* *dod*]
Q.PRS.2PL 2PL PROG know.VN **COMP.Q.be.PRS.3PL** 3PL PROG come.VN
Do you know if / whether they are coming?'
(King 1993: 305)

Some speakers introduce the conjunction *os* 'if' into the construction before the question marker (see Section 3.3 above, English influence?), compare (126) with (125) above:

(126) *Dach* *chi* *'n* *gwybod* [***os*** ***ydyn*** *nhw* *'n*
Q.PRS.2PL 2PL PROG know.VN **CONJ.if** **COMP.Q.be.PRS.3PL** 3PL PROG
dod]
come.VN
Do you know if / whether they are coming?'
(King 1993: 305)

Corresponding to *mai* in Type 1, Welsh has the interrogative *ai* which is used when the complement begins with a focused non-verbal element:

(127) *O'n* *i* *ddim* *yn* *gwybod* [***ai*** *fo* *oedd* *o*]
be.IPFV.3SG 1SG NEG PROG know.VN **COMP.Q** 3SG.M be.IPFV.3SG 3SG.M
'I didn't know if it was **him**'
(King 1993: 308)

(128) *Dos* *i* *ofyn* [***ai*** *Bert* *wed-odd* *'ny*]
GO.IMP.2SG PREP.to ask.VN **COMP.Q** Bert say-PST.3SG that
'Go and ask if it was Bert who said that'
(King 1993: 308)

Here *ai* as a complementizer is not deleted even though it generally is in main clause interrogative function:

(129) *Bert* *wedodd* *'ny*
Bert say.PST.3SG that
'Was it Bert who said that?

As these examples show, both the formal marking and semantic range of Type 2 complements in Welsh parallels that in Irish: a polar interrogative introduces complements occurring principally after main clause predicates of asking on the one hand, and those expressing some kind of doubt or uncertainty on the other.

4.3 Principles underlying the Type 1 and 2 distinctions

Clearest of all is the basic principle behind Type 2 complements: the factual status of the complement proposition is left open between two possibilities ('yes' and 'no') and neither possibility is excluded. Where other languages use conditional or indefinite pronominal markers to introduce such languages, Celtic languages use a polar question marker. Type 1 complementizers, on the other hand, simply present the complement proposition as epistemically neutral and typically carry a positive implication in respect of it. It is clear from the above that finite complementizers in Irish and Welsh form a coherent system based on their distribution: there is no possibility of combining Type 1 and Type 2, though as we have seen, harmonic combinations are possible within Type 2. When we compare Type 1-governing predicates of the kind 'know', 'be sure' with their Type 2-governing equivalents 'not know', 'be unsure', we can see the role played by polarity in determining complement type. Knowing or being sure implies a commitment to the factual status of the complement; not knowing or being unsure leaves the matter open. Relevant here are Noonan's comments (1985: 104–105) on the function of English *if* as a complementizer in canceling any positive implications of the predicate: the same function is performed in the Celtic languages by polar interrogative markers. Consider also the categories that can trigger the use of Type 2 after semifactive predicates and predicates of saying. Let us take an Irish example to illustrate this, a basic Type 1 construction as in (130):

(130) *Tá* *a* *fhios* *agam* [***go*** *bhfuil* *sé* *ann*]
be.PRS POSS.3SG.M knowing at.1SG **COMP** be.PRS.DEP 3SG.M there
'I know that he is there'

Various changes in the predicate can trigger Type 2 *an*, instead of Type 1 *go*; we find modal constructions (131), futures (132) and interrogatives (133):

(131) *Ba* *cheart* *dom* *a* *fhios* *a* *bheith* *agam* [***an***
be.COND right for.1SG POSS.3SG.M knowing PRT.to be.VN at.1SG **COMP.Q**
bhfuil *sé* *ann*]
be.PRS.DEP 3SG.M there
'I should know if he is there'

(132) *Beidh* *a* *fhios* *agam* *amárach* [***an*** *bhfuil* *sé*
be.PRS POSS.3SG.M knowing at.1SG tomorrow **COMP.Q** be.PRS.DEP 3SG.M
ann]
there
'I will know tomorrow if he is there'

(133) *An bhfuil a fhios agat* [***an*** *bhfuil sé ann*]*?*
Q be.PRS.DEP POSS.3SG.M knowing at.2SG **COMP.Q** be.PRS.DEP 3SG.M there
'Do you know if he is there?'

All of these modifications in various ways introduce uncertainty into the predicate, in the first instance, negativity (we may note however that the *coding* is consistently polar interrogative). As Frawley (1992: 390–406) points out, however, negative contexts present a scalar dimension, rather than merely a binary opposition. It is this scalar dimension that gives Type 2 complementizers their semantic coherence. This dimension can be tested simply through a negative polarity item like the Irish indefinite *aon* 'any' for each of the categories above, with the addition of conditionals. Negatives (134), interrogatives (135) and conditionals (136) pass the test of grammaticiality, while modals (137) and futures (138) do not:

(134) *Níl* ***aon*** *airgead againn*
be.PRS.NEG **any** money at.1PL
'We don't have any money'

(135) *An bhfuil* ***aon*** *airgead agat?*
Q be.PRS.DEP **any** money at.2SG
'Do you have any money?'

(136) *Má tá* ***aon*** *airgead agat*
CONJ.if be.PRS **any** money at.2SG
'If you have any money'

(137) **Ba cheart dom* ***aon*** *airgead a bheith agam*
be.COND right for.1SG **any** money to be.VN at.1SG
Intended reading: 'I should have any money'

(138) **Beidh* ***aon*** *airgead agam*
be.FUT **any** money at.1SG
Intended reading: 'I will have any money'

Uncertainty or indefiniteness underlies all of the categories above, but here we may posit a scale of indefiniteness and group negative-interrogative-conditional higher on that scale than future-modal. In Frawley's (1992) terms, within the domain of uncertainty, interrogatives and conditionals are further distinguished from futures and modals by their conceptual negativity. Another way of putting this would be to say that that negatives and interrogatives encode uncertainty while futures and modal constructions imply it (Kasper Boye p.c.). It is thus interesting to note that question and protasis markers are the sources of Type 2 complementizers in Celtic, Romance and Germanic (Nordström & Boye, this volume).

5 Complementizer omission

Welsh freely allows the omission of Type 1 *y* and Type 2 *a*, however, their original presence is reflected in the initial consonant of the following verb: compare (139) Type 1 where the verb of the complement (*bydd*) is unlenited and (140) Type 2 where it is (*fydd*):

(139) *Dw* *i* *'n* *sicr* [***bydd*** *hi* *fan hyn*]
be.PRS.1SG 1SG PRED sure **be.FUT.3SG** 3SG.F here
'I'm sure she'll be here'

(140) *Dw* *i* *ddim* *yn* *sicr* [***fydd*** *hi* *fan hyn*]
be.PRS.1SG 1SG NEG PRED sure **be.FUT.3SG** 3SG.F here
'I'm not sure if she'll be here'

By contrast, Irish strictly prohibits the omission of Type 1 *go*. Type 2 *an* may be omitted in colloquial speech especially with the verb 'to be': in this case however the verb remains in its dependent form *bhfuil*. (Contrast Germanic where only Type I complementizers may be omitted.) In general, *an* is colloquially realized as schwa; as with Welsh *a*, however, it leaves its mutation, in this case nasalization, on the following verb.

Negative complementizers cannot be omitted in either language; neither can *bod* and *i* in the superficially 'non-finite' types in Welsh.

6 Combinability issues (adverbial clauses)

We have seen above (Section 3.3) that while Type 1 and Type 2 complementizers cannot combine with each other, Type 2 complementizers can combine harmonically with each other in both Irish and Welsh. Otherwise, however, they do not combine with other subordinators. On the other hand, the more semantically neutral Type 1 *go* and *y* combine readily with other elements in Irish and Welsh respectively to form adverbial subordinators.

6.1 Irish

Most Irish adverbial conjunctions now require the nasalizing subordinator *go*. Examples presented below are causal *mar go* (141), temporal *go dtí go* (142), purposive *chun go* (143), result *ionas go* (144) and concessive *ainneoin go* and *cé go* (145 and 146). In (141), the verb form *dearnadh* is 'autonomous', meaning that

it is unmarked for subject (translates usually as an English passive). As will be apparent from the glosses, some elements in these conjunctions still function as prepositions as well:

(141) *Bhí* *fearg* *air* [***mar*** ***go*** *ndearnadh* *magadh*
be.PST anger on.3SG.M **CONJ.as** **SUB.that** make.PST.AUT mockery
faoi]
about.3SG.M
'He was angry because he had been mocked'
(based on GGBC 1960: 357)

(142) *Is* *gearr* [***go dtí*** ***go*** *dtosóidh* *siad* *air*]
be.PRS short **PREP.to** **SUB.that** start.FUT 3PL on.3SG.M
'It won't be long until they start it' (based on Ó Siadhail 1989: 324)

(143) *Éireoidh* *mé* [***chun*** ***go*** *mbeidh* *mé* *in am*]
rise.FUT 1SG **PREP.to** **SUB.that** be.FUT 1SG in time
'I'll get up early so that I'll be on time'
(based on GGBC 1960: 352)

(144) *Bhí* *sé* *crua orthu* [***ionas*** ***gur*** *imigh* *siad* *uaidh*]
be.PST 3SG.M hard on.3PL **CONJ.so** **SUB.that** leave.PST 3PL from.3SG.M
'He was hard on them with the result that they left him'
(Ó Dónaill 1977 under *ionas*, 725)

(145) ***Ainneoin*** ***gur*** *labhair* *mé* *leis*
CONJ.though **SUB.that** speak.PST 1SG to.3SG.M
'Although I spoke to him'

In the case of the copula *is*, the subordinator *go* incorporates the verb:

(146) ***Cé*** ***gur*** *láidir* *é*
CONJ.though **SUB.that.be.PRS** strong 3SG.M
'Strong as he is'
(Ó Dónaill 1977 under *cé*, 199)

6.2 Welsh

The syntactic distribution seen with Type 1 complementizers and complementation markers is transposed to the marking of adverbial clauses (all examples from King 1993: 311–319). The same syntactic situation applies in adverbial as in complement clauses: if the verb of the adverbial clause is a form of the present or imperfect of the verb 'be', the verbal noun *bod* is used (147); on the other hand, if the verb is future or conditional (148), then the appropriate inflected form is used.

In (147) the mutation from *bod* to *mod* reflects the nasalization, which would have been present historically in this construction after the 1SG possessive *fy* 'my' (here elided), something like 'so that my being able to hear'. Note however that the finite nature of the construction in present-day usage is reflected in the use of the regular subject pronoun *i* (see also 150 below):

(147) ***fel*** ***mod*** *i* *'n* *gall-u* *clyw-ed*
CONJ.SO **be.VN** 1SG PROG be_able-VN hear-VN
'so that I can hear'
(King 1993: 312)

(148) ***fel*** ***y*** *bydd-i* *di* *'n* *deall*
CONJ.SO **COMP** be.FUT-2SG 2SG PROG understand.VN
'so that you'll know'
(King 1993: 312)

In addition to purposive (final) clauses, result *fel* (149), causal *gan* (150) and concessive *er* (151) clauses are formed this way. As in complementation, the negative subordinator is *na*:

(149) ***fel*** ***na*** *ges* *i* *gyfle* *siarad*
CONJ.SO **NEG.COMP** get.PST.1SG 1SG chance speak.VN
'so that I didn't get a chance to speak'
(King 1993: 315)

(150) ***gan*** ***bo*** *chi* *'n* *mynd* *allan*
CONJ.since **be.VN** 2PL PROG go.VN out
'since you are going out'
(King 1993: 314)

(151) ***er*** ***mod*** *i* *'n* *sicr*
CONJ.though **be.VN** 1SG PRED sure
'though I'm sure'
(King 1993: 315)

As in complementation, *mai* is used if the conjunction is not directly followed by a verb. The lenition of the complement verb (*barciodd* from *parciodd*) reflects an elided relative particle (*a*):

(152) ***achos*** ***mai*** *ti* *barci-odd* *y* *car*
CONJ.because **SUB** 2SG (REL) park-PST.3SG ART car
'because it was you who parked the car' /'because **you** parked the car'
(King 1993: 314)

Most temporal conjunctions are formed by using the second superficially non-finite complement type with *i* and a verbal noun. This presumably relates to the origins of *i* as the preposition 'to' which directly mirrors the development of Irish *go*, also originally a preposition 'to', as outlined in Section 6.1:

(153) ***cyn*** ***i*** *'r* *bws* *fynd*
CONJ.before **PREP.to** ART bus go.VN
'before the bus goes'
(King 1993: 314)

(154) ***erbyn*** ***iddyn*** *nhw* *gyrraedd*
CONJ.by the time **to.3PL** 3PL arrive.VN
'by the time they had arrived'
(King 1993: 314)

(155) ***wrthi*** ***i*** *John* *ddod* *allan*
CONJ.as **PREP.to** John come.VN out
'as John came out'
(King 1993: 314)

6.3 Condition clauses

In sum, in both Irish and Welsh, the forms and structures used as Type 1 complementizers and complementation markers also play an essential role in adverbial clause structure. Significantly, in both languages, condition clauses are a conspicuous exception to this and are followed directly by the verb; (156) for Irish, (157) for Welsh:

(156) *Dá* *mbeinn*
CONJ.if be.COND.1SG
'If I were'

(157) *Pe* *byddwn* *i*
CONJ.if be.COND.1SG 1SG
'If I were'

This situation is consistent with that observed for complement clauses. Type 1 markers appear in adverbial clauses that are either presupposed or epistemically neutral;[13] whereas condition clauses resemble Type 2 complements in that they (implicitly at least) allow for two possibilities.

13 As observed for Germanic by Boye & Nordström, LANCHART Workshop, 29 October 2011. See Nordström & Boye (this volume).

7 Non-complementizing functions of complementizer forms

In both Irish and Welsh, Type 1 and Type 2 complementizers have non-complementizing functions.

7.1 Type 1

7.1.1 Irish

We will look briefly here at the use of *go* in desubordination and relativization, as well as its continuing use as a conjunction in adverbial clauses.

7.1.1.1 Desubordination

Go introduces an independent clause, expressing a wish or hope if the verb is present subjunctive; with the conditional mood the usage is exclamative (examples from Ó Dónaill 1977 under *go*, 655):

(158) ***Go*** *gcuidí* *Dia* *leo*
COMP help.SBJV God 3PL
'May God help them'
(Ó Dónaill 1977: 655)

(159) ***Go*** *mbeinn* *caillte* *murach* *iad*
COMP be.COND.1SG lost except.for 3PL
'(I declare that) I should be lost without them'
(Ó Dónaill 1977: 655)

7.1.1.2 Relativization

Irish has two types of relative clause: a direct relative, where the antecedent of the relative clause is either its subject or direct object (English type *who* or *whom*); an indirect relative, where the antecedent is indirect or prepositional object of the relative verb (English types *to whom* and *whose*). In southern dialects of Irish, an indirect relative clause is introduced by *go* (161) which replaces the relative particle *a* used in other dialects and in standard Irish (160):

(160) *an* *bhean* [***a*** *raibh* *sé* *ag* *caint* *léi*]
ART woman **REL** be.PST.DEP 3SG PROG talk.VN to.3SG.F
'the woman he was talking to'

(161) *an* *bhean* [***go*** *raibh* *sé* *ag* *caint* *léi*]
ART woman **SUB** be.PST.DEP 3SG PROG talk.VN to.3SG.F
'the woman he was talking to'

7.1.1.3 Adverbial clauses

In addition to appearing as an obligatory component of compound adverbial conjunctions, *go* can also appear as a free standing conjunction in a variety of adverbial clauses, temporal (162), purposive (163), result (164) and causal (165). The examples are from Ó Siadhail (1989: 324–325):

(162) *Is* *gearr* [***go*** *dtosóidh* *siad* *air*]
be.PRS short **CONJ** start.FUT 3PL on.3SG
'It won't be long until they start it'
(Ó Siadhail 1989: 324–325)

(163) *Teanam* [***go*** *raibh* *deoch* *againn*]
go.IMP.1PL **CONJ** be.SBJV drink at.1PL
'Let's go (so that we can) have a drink'
(Ó Siadhail 1989: 324–325)

(164) *Bhí* *an* *oíche* *chomh* *dubh* [***go*** *ndeachaidh* *mé* *treasna* *ar*
be.PST ART night so dark **CONJ** go.PST.DEP 1SG across PREP.on
an *gclaidhe*]
ART ditch
'The night was so black that I went over the stone wall'
(Ó Siadhail 1989: 324–325)

(165) *Chuir* *sé* *fios* *ar* *mo* *sheanathair* *a* *thíocht* *aige*
PST\put 3SG knowing on POSS.1SG grandfather PRT.to come.VN to.3SG
[***go*** *raibh* *gnatha* *aige* *dó*]
CONJ be.PST.DEP business at.3SG.M for.3SG.M
'He sent for my grandfather to come to him since he required him'
(Ó Siadhail 1989: 324–325)

Go can also be used in clause coordination, where it is virtually synonymous with 'and' (166); it also functions as a preposition 'to' (167):

(166) *Tháinig* *sé* *isteach* *sa* *seomra* [***go*** *bhfuair* *sé*
come.PST 3SG into PREP.in.ART room **CONJ** get.PST 3SG
an *leabhar*]
ART book
'He came into the room and got the book'
(Ó Dónaill 1977 under *go*, 655)

Go also retains its original function as a preposition (167):

(167) ***go*** *hÉirinn*
PREP.to Ireland
'to Ireland'

7.1.2 Welsh

Outside complement structures, only in Literary Welsh is the particle *y* found. As we have seen (Section 4.1.2 above), it is freely omitted in complementation.

7.1.2.1 Main clause *y*

Y used as an affirmative preverbal particle before the present tense form of *bod* 'be' (*mae*):

(168) ***Y*** *mae* *Gwen* *yn* *yr* *ardd*
AFF be.PRS.3SG Gwen PREP.in ART garden
'Gwen is in the garden'
(Borsley et al. 2007: 35)

The particle *y* appears here as a general affirmative marker with no particular semantic content; in spoken Welsh it is simply omitted.

7.1.2.2 Relativizer *y*

In Welsh, as in Irish, wh-questions have the structure of relative clauses. The particle *y* appears in relativizer position:

(169) *Gan* *bwy* [***y*** *cefaist* *y* *llythyr* *hwnnw*]*?*
PREP.from Q.who **REL** get.PST.2SG ART letter that
'Who did you get that letter from?'
(Borsley et al. 2007: 115)

Otherwise *y* occurs as a relative marker in indirect relatives where the antecedent is the object of a preposition (170) and possessor relatives (171) equivalent to English 'whose' (compare the similar use of Irish *go* in Section 6.1.2.2 above):

(170) *y* *wraig* [***y*** *gwerth-odd* *Ieuan* *y* *ceffyl* *iddi*]
ART woman **REL** sell.PST-3SG Ieuan ART horse to.3SG.F
'the woman that Ieuan sold the horse to'
(Borsley et al. 2007: 121)

(171) *y* *dyn* [***y*** *gwel-ais* *ei* *chwaer*]
ART man **REL** see.PST-1SG POSS.3SG.M sister
'the man whose sister I saw'
(Borsley et al. 2007: 122)

7.1.2.3 Particle *i*

The particle *i* otherwise functions as the preposition 'to' (172) and can also be used as a final conjunction (173):

(172) *Dych* *chi* *'n* *mynd* [***i*** *'r* *dre* *heddiw?*]
Q.be.PRS.2PL 2PL PROG go.VN **PREP.to** ART town today
'Are you going to town today?'
(King 1993: 281)

(173) ***i*** *ddal* *y* *bws*
PREP.to get.VN ART bus
'(in order) to catch the bus'
(King 1993: 282)

7.1.2.4 *Bod*

As we have already seen, the complement marker *bod* is the verbal noun of the verb 'be'. See the discussion at various points above for the arguments that in its use in complementation it is essentially finite from a syntactic point of view (see for example Section 4.1.3.2).

7.2 Type 2

7.2.1 Polar questions

The Type 2 affirmative complementizer in both Irish and Welsh is the polar question marker, Irish *an* (174) and Welsh *a* (175; see Section 1 above and *passim*):

(174) ***An*** *bhfeic-eann* *tú* *í?*
Q see-PRS 2SG 3SG.F
'Do you see her?'

(175) ***A*** *gaeth* *Ffred* *wobr?*
Q get.PST.3SG Fred prize
'Did Fred win a prize?'

7.2.2 Irish condition markers

The Irish protasis markers *dá* (Section 3.2.3) and *má* (Section 3.2.5) 'if' are marginal to the complementation system, as we have seen. Nevertheless, here are some examples of their use outside complementation. Both are used as protasis markers, *má* in past, present and future conditionals (176), *dá* in hypothetical and counterfactual ones (177). In addition, both conjunctions can be used exclamatively ([178] and [179]):

(176) [***Má*** *dhéan-ann* *tú* *sin*], *beidh* *tú* *thíos* *leis*
CONJ.if do-PRS 2SG that be.FUT 2SG down with.3SG.M
'If you do that, you'll regret it'

(177) [***Dá*** *ndéan-fá* *sin*], *bheifeá* *thíos* *leis*
CONJ.if do-COND.2SG that be.COND.2SG down with.3SG.M
'If you did that, you'd regret it'

(178) ***Má*** *chonaic* *féin!*
CONJ.if see.PST even
'Even if (you) saw (it)!'

(179) ***Dá*** *bhfeic-feá* *é!*
CONJ.if see-COND.2SG it
'If (only) you had seen it!'

Finally, we can compare the use of *má* in focus constructions marking two opposed values (180) with the use of *cé acu* in Type 2 complementation (181). In (180), the form *más* combines the conjunction *má* with the present copula *is*:

(180) *Más* *te* *más* *fuar* *an* *lá*
CONJ.if.be.PRS hot CONJ.if.be.PRS cold ART day
'Let the day be hot or cold'
(Ó Dónaill 1977 under *má*, 813)

(181) *Is* *cuma* *liom* [***cé*** ***acu*** *a* *bhí* *sé* *te* *nó* *fuar*]
be.PRS same to.1SG **which** **of.3PL** REL be.PST 3SG.M hot or cold
'I don't care whether it was hot or cold'

Both structures allow for two possibilities, or two polar extremes.

7.3 Irish *nach* and Welsh *na*

In both Irish and Welsh, the negative complementizer, Irish *nach* (past *nár*) and Welsh *na* (*nad* before a vowel), is in fact a general negative subordinator that appears also in adverbial and relative clauses of the kinds discussed above.

7.4 Summary

Analysis of the non-complementizing functions of the Irish and Welsh particles is again consonant with their behavior in complementation. Irish *go* appears in adverbial clauses; both *go* and Welsh *y* (with restrictions of dialect and register respectively) show up in relative clauses. Once again, these are contexts which are generally presupposed or epistemically neutral. On the other hand, the Type 2 complementizers of both languages are polar interrogatives, which by their very nature entail the non-exclusion of alternative scenarios ('yes' or 'no').[14]

8 Diachrony

As was stated above (Section 1.1), a notable feature of Celtic complementizers is that the negative forms have no relation to the affirmative ones. In this section we will examine the diachronic sources of the complementizers or complementation markers under investigation, and the development of their semantic functions where relevant.

8.1 Type 1

8.1.1 Irish

Here we will examine briefly the origins of *go* and *nach* before taking a more detailed look at the role of *go* in the historical development of finite complementation.

14 In more or less the terms in which Fagard, Glikman and Pietrandrea discussed the Romance evidence, LANCHART Workshop, October 2011. See further this volume.

8.1.1.1 Origins of *go*

Type 1 *go* derives from the Old Irish preposition *co* 'to, as far as' (McCone 1993). Already in Old Irish (by the eighth century) we find a conjunction *co* with temporal ('until'), consecutive ('so that') and purposive ('in order that') meanings and we have the beginnings of *co* as complementizer after verbs of saying. The conjunction consists of the preposition augmented by an indirect relative particle *a* (from a preform **san*), explaining the nasal mutation on the following verb (Thurneysen 1980: 298–299). The original semantics of *co* are allative and we might conceive of its development in a very rudimentary way as follows (see Traugott 1978):

A co B = A as far as/to/till B (spatial and temporal)
> A causes B so B follows A (i.e. *post hoc ergo propter hoc*)
> A (generally) governs B (discourse)

As we have seen above, most adverbial clauses now require the presence of this formal subordinator *go*.

8.1.1.2 Origins of *nach*

In Old Irish the form of the negative complementizer is *nád* (as in [182]) which is in origin the negative relative marker (183):

(182) *dutluchtetar* *ní* [***nád*** *tardatis*]
desire.PRS.3PL thing **NEG.REL** give.PST.SBJV.3PL
'They desire something which they had not given'
(McQuillan 2002: 203)

(183) *asbiur-sa* [***nád*** *ningén*]
say.1SG-EMPH **NEG.COMP** know.PRS.1SG
'I say that I know not'
(Ó hUiginn 1999: 123)

In present-day southern Irish dialects, this original form *nád* is continued as the negative complementizer *ná* causing no initial mutation rather than the standard form *nach*^{N} which has featured in the discussion thus far:

(184) *Deir-eann* *sí* [***ná*** *beidh* *sí* *ann*]
say-PRS 3SG.F **NEG.COMP** be.FUT 3SG.F there
'She says that she won't be there'

The modern standard form *nach* consists of Old Irish *nád* (*nad*) with a fossilized neuter infixed object pronoun and is generalized in most forms of the language as the negative subordinator during the Middle Irish period.

8.1.1.3 Origins of *go* as a complementizer

In Section 8.1.1.2 we have seen how negative complements are marked in Old Irish and how they originate in relative clauses. Let us look now at affirmative complements.[15] There are two options: parataxis and relativization. The original situation is represented by (186) in which the original clause of (185) is simply juxtaposed with its 'main' clause:

(185) *Attá* *lat-su*
be.PRS.3SG with.2SG-EMPH
(it is with you)
'You have it'

(186) *Is* *derb* *limm* *attá* *lat-su*
be.PRS certain with.1SG be.PRS.3SG with.2SG-EMPH
'I am certain you have it'
(Ó hUiginn 1999: 123)

An option that becomes more common in the later Old Irish period is a strategy borrowed from relative clauses: the use of a nasalizing relative (in 187 the complement verb form is relative *bías* 'that is', here nasalized as *mbías*):

(187) *arna* *ro-chretea* [***mbías*** *ícc* *do*
so that.NEG PRT-believe.SBJV.3SG **be.PRS.HAB.REL** salvation for.3SG.M
hua *dia*]
PREP.from God
'that he may not believe that he shall have salvation from God'
(Ó hUiginn 1999: 123)

The subjunctive mood appears in Early Irish complements, in the oldest example here (188) we have parataxis; in (189) the subjunctive is preceded by *ara*, otherwise a final conjunction, while in the chronologically latest example of the three, the complementizer is *co* (190), which also appears in adverbial purpose clauses:

15 For a discussion of complement structures and strategies in Early Irish with verbs of saying, see Ó hUiginn Í1999); for the subjunctive in Early Irish complement clauses, see McQuillan (1999; 2002: 55–10).

(188) *Gádatar* ***tíssed*** *in* *nóeb*
pray.PST.3PL **come.PST.SBJV.3SG** ART saint
'They prayed (that) the saint might come'
(McQuillan 2002: 61)

(189) *Gigeste-si* *dia* [***ara*** *fulsam*]
pray.FUT.2PL-EMPH God **COMP** endure.SBJV.1PL
'You will pray God that we may endure'
(McQuillan 2002: 64)

(190) *Guid-im* [***co*** *ndeach-air-si*]
pray-PRS.1SG **COMP** go.SBJV-SG-EMPH
'I pray that you (may/should) go'
(McQuillan 2002: 47)

The locus for the rise of discrete complementizers in Early Irish is the subjunctive complement. This is because the subjunctive becomes an increasingly syntacticized category dependent for its usage on the presence of other markers, typically specific conjunctions. Two of the most common of these conjunctions are those used above as complementizers in (189) and (190):

– Purposive *ara* 'in order that'
– Purposive/Consecutive/Temporal) *co* 'in order that', 'so that; 'until'

What we find is that the use of *co* and *ara* in complementation is harmonic with their use in adverbial clauses of purpose. In (189) for example, *ara* could easily be taken as introducing a final clause: 'You will pray God in order that we may endure', similarly (190): 'I pray so that you may go'. In each case, the speaker in the main clause wishes to 'manipulate' the event of the subordinate clause.

My thesis here for the rise of *co* as a general affirmative complementizer rests initially on its association with the subjunctive. The increasingly syntacticized nature of the subjunctive mood provides the initial impetus for the formal marking of finite complement clauses in Irish by means of conjunctions. As part of its overall desemanticization in Old Irish, the subjunctive becomes marginal in precisely those structures used in complementation: main clauses used paratactically and relative clauses. Consider the following pair of examples for main clause usage. The Old Irish subjunctive is used independently as an optative to express a wish or a hope (191 *rom-bet*). However, already in Old Irish the conjunction *co* appears in this main clause function before the subjunctive and eventually becomes obligatory (192 *condom-roib*):

(191) *Ro-* *<m>* *bet*
PRT <OBJ.PRON.1SG> be.SBJV.3PL
(may they be to me)
'May I have them'
(McQuillan 2002: 45)

(192) ***Co*** *<ndom>* *roib*
COMP <OBJ.PRON.1SG> be.SBJV.DEP.3SG
(that it may be to me)
'May I have it'
(McQuillan 2002: 46)

Another aspect of this syntacticization is that increasingly in relative structures, the subjunctive is confined to clauses where the antecedent is both indefinite and under negative or irrealis scope as in (193):

(193) *Ní* *bía* *nech* [*ro -n -íccae-su*]
NEG be.FUT.3SG anyone PRT-REL-save.SBJV-EMPH
'There will be no one who can save you'
(McQuillan 2002: 204)

Accordingly, the conjunctions *co* and *ara* as used in adverbial clauses, especially purpose or final clauses, are increasingly pressed into service as complementizers in affirmative clauses when the verb is subjunctive because the subjunctive can no longer stand independently as a verbal category.

Due to extensive changes in the Middle Irish verbal system as a whole, the nasalizing relative collapses in the tenth century, however the form *nach,* which originates as a negative relative marker endures as the regular negative complementizer (Ó hUiginn 1986). By the twelfth century, *co* is on its way to becoming the affirmative finite complementizer, both in subjunctive and non-subjunctive clauses, although paratactic structures do endure for a while longer. Initially, we see some development of *ara* as a genuine complementizer, that is to say, in contexts divorced from that of purpose, as in following the main predicate 'know not':

(194) *Ni* *fitir* [***ara*** *mbeth* *and*]
NEG know.PRS.3SG **COMP** be.PST.SBJV.3SG there
'He did not know that it was there'
(McQuillan 2002: 77)

Ultimately, however, what militates against the development of *ara* as a complementizer is its confinement to subjunctive contexts. In other words, while subjunctive clauses provide the initial impetus for complement-marking by conjunctions, it is the greater versatility of *co* in combining with indicative and conditional as well as subjunctive categories that gives it an advantage over *ara*.

The key to this is the idea of *consecutiveness*: the syntagm *A co B* establishes a flexible relationship between the two clauses within the parameters: 'A as far as/followed by B' where the semantic nature of A determines the precise relationship between them. Compare the range of the two conjunctions in Early Irish clauses (Thurneysen 1980: 554–556):

Table 3: Range of *ara* and *co*

Ara	**Verb of *Ara*-Clause**
Purposive	Subjunctive
Co	**Verb of *Co*-Clause**
Temporal 'until'	Indicative/Subjunctive
Consecutive 'so that'	Indicative (Result/Narrative) Subjunctive (Inference)
Purposive	Subjunctive
Complement	Indicative/Subjunctive

Historically, *co* as a complementizer then extends to indicative and conditional contexts, as in the following twelfth-century example. Note in (195) how the complementizer incorporates the preverbal particle *ro* preceding the verb (*ghonsat*):

(195) *Ro-fetamair* [***co-r*** *ghonsat* *ar* *n-aithre* *slis*
PRT-know.PRS.1PL **COMP-PRT** wound.PST.3PL POSS.1PL father.PL side
7 *toeb* *Ísu*]
and flank Jesus
'We know that our fathers wounded the side of Jesus'
(McQuillan 2002: 85)

From Table 3, the so-called narrative use of *co* merits special mention. In this usage, *co* with the past indicative simply introduces a series of actions or events in sequence, each one subsequent to the last. In the following, each finite verb form is preceded by *co*:

(196) ***co*** *mben-som* ***co*** *rruc* *tulchlár* *ind*
CONJ strike.PRS.3SG-EMPH **CONJ** drive.PST.3SG front ART.GEN
étain ***co*** *mboí* *forsind* *innchinn* ***co-r***
forehead.GEN **CONJ** be.PST.3SG on.ART brain **CONJ-PRT**
thrascair
knock down.PST.3SG
'and with his fist he struck him on the forehead, driving the front of his forehead onto his brain, while with his arm he knocked (it) down'
(McQuillan 2002: 127)

Here *co* has the force of a simple connective ('and then [verb] ... and then [verb]'; see [166] above for a modern example). In any case, by the Early Modern Irish period, *co*, later *go*, has firmly established itself as the affirmative finite complementizer. I would argue that this is on the strength of its versatility and semantic neutrality: it can be used with all finite verb categories that appear in subordinate clauses while it can simply have the semantic specificity of 'and'. Thus in the course of the Middle Irish period, *co* emerges as the particle ideally suited to mark a range of epistemically neutral subordinate clauses.

8.1.2 Welsh

All the complement structures described above for Modern Welsh are already found in the Middle Welsh period which runs from the twelfth to the end of the fourteenth century.

8.1.2.1 Origins of *y*

As we saw in Section 2.3 above, Welsh is rich in preverbal particles. Historically, the complementizer *y* continues the most common preverbal affirmative particle found in Middle Welsh. It is found in both main (197) and complement (198) clauses. The usage represented by (197) is continued in Literary Welsh in the form *y mae* (see Section 7.1.3.1):

(197) ***Y*** *kyuodes* *y* *marchawc*
AFF rise.PST.3SG ART knight
'The knight arose'
(Evans 1976: 171)

(198) *Gwybydwch* *chwi* [***y*** *daw* *kennad-eu*]
know.IMP.2PL 2PL **COMP** come.FUT.3SG messenger-PL
'Know ye that messengers will come'
(Evans 1976: 171)

We have seen above that various relative clause uses of *y* are found in the modern literary language (Section 7.1.3.2 above). These are also attested in Middle Welsh, for example:

(199) *Y* *brenhin* [***y*** *kiglef* *y* *glot*]
ART king **COMP** hear.PST.1SG POSS.3SG.M fame
'The king whose fame I have heard of'
(Evans 1976: 65)

Literally (199) is something like: 'the king, I have heard of his fame'. A morphophonological characteristic of *y* at all periods is that it does not mutate the following verb.

8.1.2.2 Origins of *na(d)*

The modern finite negative *na(d)* continues Middle Welsh *na(t)* (*t* = /*d*/). In both the medieval and modern language, these forms function as general subordinate negators. In main declarative finite clauses in Middle Welsh, the negative is *ny(t)*; in spoken Modern Welsh, this has been ousted by the adverbial *ddim* 'anything, at all' which was originally used to emphasize the negation but has now become the primary, obligatory, negative marker in main clauses. The form *ddim* has also spread to complement clauses but remains optional here in some cases (see Section 4.1.2.2 and especially examples [74] and [75]).

8.1.2.3 Origins of *mai*

In origin, *mai* as a complementizer is historically the same as *mae*, 3rd singular present indicative form of *bod* 'to be'. In modern Literary Welsh, *mae* is preceded by the particle *y*, see (168) above; already in Middle Welsh, *y mae* is used to introduce complement clauses where the first element is a noun or adjective ('that it is'):

(200) *A* *thi* *a* *wybydy* [***y*** ***mae*** *gwac* *uocsach*]
And 2SG PRT know.FUT.2SG **AFF** **be.PRS.3SG** boast vain
'And thou shalt know that (it) is a vain boast'
(Evans 1976: 144)

It is from instances such as this that modern *mai* evolves.

8.1.2.4 Origins of 'non-finite' complement marking

In Middle Welsh, the finite-*bod* construction is more widespread than in the modern language, in that any affirmative verb may appear as a verbal noun complement to Type 1 predicates (propositional attitude, utterance, perception) as long as the complement event is anterior to the predicate:

(201) *mi* *a* *gigleu* [***dyuot*** *y* *'r* *Deheu* *y ryw* *bryuet*]
1SG PRT hear.PST.3SG **come.VN** to ART south the sort creatures
'I have heard that creatures the like of which ... have come to the south ...'
(Borsley *at al*. 2007: 327)

In the course of the modern period, the range of this construction contracts, so that it now affects only the present and imperfect forms of the verb *bod* in the complement clause.

The range of the other formally non-finite construction with *i* + complement subject + verbal noun has expanded. In Middle Welsh the preposition *i* 'to' has the form *y*:

(202) *Y* *dywedant* *wy* [***y*** Adaf gwynaw Abel]
AFF say.PRS.3PL 3PL **PREP.to** Adam lament.VN Abel
'They say that Adam mourned for Abel'
(Evans 1976: 162)

(203) *ny* *thebygaf* *i* [***y*** *un* *o* *hyn* *uynet*]
NEG think.PRS.1SG 1SG **PREP.to** one of them go.VN
'I do not think that any of these people will go'
(Borsley *at al*. 2007: 333)

In Middle Welsh, the construction has not yet become confined to events anterior to the main clause, as in Modern Welsh. According to Evans (1975: 162), this construction is rare in Middle Welsh but becomes the most common means in the modern language of expressing past time in complement and adverbial clauses. Building on the analysis of Richards (1949–51), Borsley et al. (2007) suggest that the origin of this construction lies in the reanalysis of complements to raising verbs such as *daruot* 'happen, finish' and *damwein(y)aw* 'happen'. This latter verb can be followed by either superficially non-finite complement type: the *bod*-type as in (204) or preposition *i* (medieval *y*) followed by verbal noun as in (205) (Borsley et al. 2007: 332):

(204) *y* *damweinawd* [***dyuot*** *llu* *o* *'r* *Flemesseit*]
AFF happen.PST.3SG **come.VN** force PREP.of ART Fleming.PL
'It happened that a force of the Flemings came'

(205) *ef* *a* *damweinyawd* [***y*** ***wynt*** ***dwyn*** *yr* *ysgraff* *ymeith*]
3SG.M PRT happen.PST.3SG **PREP.to** **wind** **take.VN** ART boat away
'It happened that a wind took the boat away'

In (205) at some point, the phrase *y wynt* 'the wind' moves from being analyzed as the indirect object of the main clause to being the embedded subject of the complement. Once that happens, the construction spreads to all verbs taking Type 1 complements which already have construction (208) above. The construction simultaneously extends to adverbial clauses (especially temporal) in the early modern period (see Section 6.2 above).

8.2 Type 2

8.2.1 Irish

Irish Type 2 *an* continues the form and function of the Old Irish interrogative *in* (206) which already in the older language appears in indirect questions (207). Note the form *im* before a labial (Thurneysen 1980: 291):

(206) *In biam?*
Q be.FUT.1PL
'Will we be?'

(207) *As-rubart i(m) mboí*
PRT-PRT.say.PST.3SG Q be.PST.3SG
'He asked whether there was'

A well-attested construction in Old Irish is for the question marker to be preceded by *do fhius(s)* 'to know' (see the use of Modern Irish *a fhios* 'knowing, knowledge' as a complement-taking predicate in Section 4.1.2 above):

(208) *Fo bíith precept dóib* ***dúus*** ***in*** *-duccatar*
because preaching to.3PL **PREP.to_know.VN** **Q** bring.SBJV.3PL.PASS
fo hiris
under faith
'because of preaching to them to know if they may be brought under the faith'

8.2.2 Welsh

Type 2 *a* continues the same form from the medieval period as the direct polar question marker (209); *ae* (precursor of modern *ai*) is found introducing questions where the verb is not the first element (210):

(209) ***A*** *uyd llawn dy got ti uyth?*
Q be.FUT.3SG full POSS.2SG bag 2SG ever
'Will thy bag ever be full?'
(Evans 1976: 174)

(210) ***Ae*** *ti* [*a eirch uy merch?*]
Q 2SG REL seek.PRS.2SG POSS.1SG daughter
'It is thou who seekest my daughter?'
(Evans 1976: 175)

The same markers are found in the corresponding indirect questions:

(211) *yn* *gofyn* *y* *Idawc* [***a*** *gaffei* *ran*]
PROG ask.VN PREP.to Iddawg **Q** get.COND.3SG share
'asking Iddawg if he might have a share'
(Evans 1976: 174)

(212) *Gouynnwch* *idi* [***ae*** *ryd*]
ask.IMP.2PL to.3SG.F **Q** free
'Ask her whether she is free'
(Evans 1976: 175)

The Welsh system has remained essentially stable over time; however, as we have seen, the various particles are much more likely to be omitted in popular speech today.

9 Summary

Despite considerable differences in detail, as representatives of the two subbranches of the modern Celtic languages, Irish (Goidelic) and Welsh (Brittonic) show a fundamental semantic similarity in the organization of their finite complement structures. A basic distinction is made in both languages between a set of epistemically neutral complementizers on the one hand (Type 1), and a set expressing epistemic uncertainty on the other (Type 2). Type 1 present the complement proposition as the one and only possibility, Type 2 leave the status of the proposition open. The basic two-fold semantic distinction of Celtic is paralleled in both Romance and Germanic (Nordström & Boye, this volume). The behavior of each set of complementizers outside complementation supports this conclusion: Type 1 occur in contexts that are neutral or presupposed (adverbial and relative clauses), Type 2 are polar question markers. The diachronic origins of the markers are also consistent with the analysis. As far back as our records of each language go, Type 2 have always been interrogatives; Type 1 on the other hand are more disparate in origin. Welsh *y* originally marks a clause as affirmative while Irish *go*, originally a preposition, conveys a basic sense of consecutiveness or sequentiality.

Acknowledgment

My special thanks go to Professor John Koch, University of Wales, Aberystwyth, for his help with the Welsh material.

References

Ball, Martin J. & James Fife (eds.). 1993. *The Celtic Languages*. London and New York: Routledge

Borsley Robert D., Maggie Tallerman and David Willis. 2007. *The Syntax of Welsh*. Cambridge: Cambridge University Press.

deBhaldraithe, Tomás. 1956/1957. Aimhreas; mar a chuirtear i gcaint é [*aimhreas* as used in speech]. *Éigse* 8. 147–149.

Evans, D. Simon. 1976. *A Grammar of Middle Welsh*. Dublin: Dublin Institute for Advanced Studies.

Frawley, William. 1992. *Linguistic Semantics*. New Jersey: Lawrence Erlbaum Associates.

GGBC. 1960.*Graiméar Gaeilge na mBráithre Críostaí* [Christian Brothers' Irish grammar]. Dublin: Mac an Ghoill.

Givón, Talmy. 1995. *Functionalism and Grammar*. Amsterdam: John Benjamins.

King, Gareth. 1993. *Modern Welsh. A Comprehensive Grammar*. London and New York: Routledge.

McCone, Kim. 1993. Old Irish *co*, *cucci* 'as far as (him, it)' and Latin *usque* 'as far as'. *Ériu* 44. 171–176.

McQuillan, Peter. 2002. *Modality and Grammar. A History of the Irish Subjunctive*. Maynooth: Department of Old Irish.

McQuillan, Peter. 1999. Complementation and the subjunctive in early Irish. *Ériu* 50. 87–132.

Noonan, Michael. 1985. Complementation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. 2, *Complex constructions*, 42–140. Cambridge: Cambridge University Press.

Nordström, Jackie & Kasper Boye. This volume. Complementizer semantics in the Germanic languages.

Ó Dónaill, Niall (ed.). 1977. *Foclóir Gaeilge-Béarla*. Dublin: Oifig an tSoláthair.

Ó hUiginn, Ruairí. 1986. The Old Irish nasalizing relative clause. *Ériu* 37. 69–75.

Ó hUiginn, Ruairí. 1998. Complementation and the *verba dicendi*. *Ériu* 49. 121–148.

Ó Siadhail, Mícheál. 1989. *Modern Irish. Grammatical Structure and Dialectal Variation*. Cambridge: Cambridge University Press.

Quinn, E. G. (ed.). 1983. *The Dictionary of the Irish Language Compact Edition*. Dublin: Royal Irish Academy.

Rhys Jones, T. J. 1977. *Living Welsh*. New York: Hodder and Stoughton.

Thurneysen, Rudolf. 1980 [1946]. *A Grammar of Old Irish.* Dublin: Dublin Institute for Advanced Studies.

Traugott, Elizabeth. 1978. On the expression of spatio-temporal relations in language. In J.H. Greenberg (ed.), *Universals of Human Language*, vol. 1, 369–400. Stanford: Stanford University Press.

Traugott, Elizabeth. 1989. On the rise of epistemic meaning: an example of subjectification in language change. *Language* 65. 31–35.

Benjamin Fagard, Paola Pietrandrea and Julie Glikman

Syntactic and semantic aspects of Romance complementizers

1 Introduction

This article aims at describing formal and functional properties of Romance complementizers. The paper is organized as follows: we provide an overview of Romance languages (sections 1.1. and 1.2.) as well as of the methodology employed in our study (Section 1.3.). We propose functional and formal definitions of complementation, complementizers and 'canonical' complementizers (Section 2.1.); we briefly sketch an overview of Romance complementizers and we justify our focusing on the Romance complementizers corresponding to the English *if* and *that* (Section 2.2.). We describe the syntactic behavior of these complementizers (sections 2.3. and 2.4.). We then focus on their semantic aspects and distribution; we show that a general function of illocutionary subordination is common to all complementizers, as is proven by the semantic effects of complementizer omission (Section 3). We study the semantic distributin of Romance 'that' complementizers and of Romance 'if' complementizers (Section 4.1). A semantic analysis of these complementizers then enables us to specify which type of non-factuality is encoded in complements introduced by 'if' complementizers (Section 4.2 and 4.3), as well as in complements introduced by 'that' complementizers when they are in the indicative and when they are in the subjunctive (4.4). We focus on the marginal deviations found in Romanian on the one hand and French on the other hand as compared to other Romance languages (5). We finally focus on the diachronic sources of Romance complementizers (6).

1.1 Romance languages – sociolinguistic and geographical features

Romance languages are a subgroup of Indo-European languages originating from Latin. They are mostly spoken in Europe (where all Romance languages are attested), in the Americas (mainly French, Spanish and Portuguese), and in Africa (mainly French and Portuguese), but there are also smaller, isolated communities elsewhere. The total number of mother-tongue speakers is around 700 million,

among which around 200 million in Europe (by Ethnologue[1] counts, which we compiled). There are also Creoles based on Romance languages, with a total of around 11 million speakers. For instance, Creole languages based on French are found in Martinique, Guadeloupe, Haiti, French Guyana, Réunion and Louisiana, with close to 10 million speakers in total (according to the Ethnologue website, these are the following varieties: Saint Lucian, Guianese, Haitian, Karipúna Creole French, Louisiana Creole French, Guadeloupean, Morisyen, Réunion Creole French, San Miguel Creole French, Seselwa and Tayo). Chavacano and Palenquero are Spanish-based Creoles and have close to 300,000 speakers in total. The Portuguese-based Creoles are Angolar, Cafundo Creole, Crioulo, Fa D'ambu, Indo-Portuguese, Kabuverdianu, Korlai Creole Portuguese, Macanese, Malaccan Creole Portuguese, an extinct Pidgin in East Timor, Principense, Sãotomense, and Ternateño (with close to 700,000 speakers in total). However, from a grammatical point of view at least –which is our concern here– Creoles in general, and even Romance-based ones, are very different from European Romance, and it does not seem justified to address them jointly: we believe Romance-based Creoles deserve a separate, dedicated study. For questions of simplicity, we shall henceforth use the term "Romance" to refer exclusively to European Romance languages, to the exclusion of Romance-based Creoles. Besides, although Romance languages proper are also spoken in South America, Africa and marginally elsewhere, our study focuses on European varieties.

Table 1: Number of speakers for the main Romance languages (in millions)

Variety	Number of speakers in the country of origin	Total number of speakers
Romanian	19,7	23,4
Romansch	0,035	0,035
Italian	55	61,7
French	53,2	67,8
Catalan	11,2	11,5
Occitan	1,9	2
Spanish	28,2	328,5
Portuguese	10	178
Galician	3,2	3,2

1 In its web version (www.ethnologue.com). Note that some numbers seem low; for French, for instance, by the 2010 counts of the 'Organisation internationale de la francophonie', the actual number of native speakers in the world is 115 million. The numbers in Table 1 should therefore be taken with caution; however, they do provide a rough idea of the number of speakers and the relative (numerical) importance of each Romance language.

1.2 Romance languages – major typological characteristics

The major typological characteristics of Romance languages include the fact that most of them are "pro-drop" languages, i.e. that subject pronouns tend to be omitted, or can be omitted. Note that this is not the case of French, nor of some Romance varieties in Northern Italy and Switzerland. A partly correlated feature of Romance is its relatively free word order, in pro-drop Romance languages. Romance languages are by and large inflectional as far as verbs are concerned, with a heavy Tense-Aspect-Modality system, but not so much for nouns and adjectives, which display variable genre and number marking and no case marking, with the exception of Romanian. Romance languages generally have preposed definite and indefinite articles, a large set of prepositions and conjunctions, and few modal particles comparable to the German modal particles such as *ja, auch, denn, doch* (cf. Diewald and Kresić 2010; Squartini 2013; Schoonjans 2013), relying instead on discourse markers.[2] Finally, they have infinitive clauses and a synthetic future tense. French differs from "Central Romance" on a number of points besides word order and pronouns; it displays for instance an analytic future tense and two distinct paradigms for demonstratives (demonstrative pronouns vs. demonstrative determiners). Romanian, on the other hand, presents a set of distinctive Balkan features. These include, for the verbal sphere, an analytic future tense (actually, two: [*o să* + subjunctive], based on the complementizer *să*, and [*voi* + infinitive], based on the modal verb *a vrea* 'to want'), a wide use of the subjunctive mood and (correlated) avoidance of the infinitive. Romanian also has doubled clitics (e.g., *am văzut-o pe Maria* 'I saw (her) Maria'). For the nominal sphere, the most notable features are the postposed definite article (e.g. *lupul* 'the wolfe.NOM/ACC') and a residual set of nominal case markers with genitive-dative fusion (e.g. *lupule* 'the wolfe.VOC', *lupului* 'the wolfe.GEN/DAT'). Finally, the lexicon has an important Slavic component absent from other Romance languages. This specificity of French on the one hand (see e.g. Koch 2002) and Romanian on the other can be observed in the case of complementizers, too, as we will see in this paper.

2 For the distinction between modal particles and discourse markers, we refer to Degand, Cornillie & Pietrandrea (2013). The authors define modal particles as "a class of distributionally constrained words dedicated to semantically qualify the speech act with regard to the pragmatic presupposed context" and they define discourse markers as "linguistic elements endowed with the indexical function of relating the utterance in their scope to the linguistic context" (Degand, Cornillie, and Pietrandrea 2013: 23–24, but see also Diewald 2013 for a detailed discussion).

1.3 Methodology

In the following sections we investigate formal and functional aspects of Romance complementizers. In order to investigate their uses, we combined different methodologies. In a first phase, we based our hypotheses on the existing literature (for instance Jeanjaquet 1894; Herman 1963; Bat-Zeev Shyldkrot & Kemmer 1988), and on traditional philological instruments, i.e. dictionaries and grammars (see the Reference section). In a second phase, we called on native speakers' judgments to confirm these claims. In a third phase, we relied on corpora to check the validity of our results. The use of corpora for determining how speakers make use of language is now quite well established, as one can see from the flourishing subfield of corpus-based grammars (e.g., for English: Biber et al. 1999; for French: Abeillé, Godard & Delaveau in press, and in diachrony Marchello-Nizia et al. in prep. etc.). It is based on the idea that linguistic systems are best viewed as language(s) in use, and that grammar is shaped by frequency effects (see e.g., Bybee &Hopper 2001) – in other words, it seems to us that it is more interesting to look at language use and find out from there what rules the speakers are following than to rely exclusively on introspection. The corpora we used are the following (links to the corresponding databases are given in the Reference section):

- French: Valibel, PFC, Clapi, Frantext, frTenTen;
- Italian: Badip, itTenTen;
- Portuguese: Corpus do Português (Davies & Ferreira 2006), ptTenTen11;
- Romanian: CORV (Dascălu-Jinga 2002), Ruxândoiu, Romanian web corpus;
- Spanish: CREA, CORDE, Corpus del Español (Davies 2002), esTenTen.

We believe that these 'cross-examinations' guarantee the validity of our results[3].

3 We would like to thank here the friends and colleagues who accepted to collaborate, especially Alexandru Mardale, José Carlos de Hoyos, Lennys Ospino, Daniel Rojas-Plata, Ana Fonseca, Sonia Pereira, Ana Margarida Abrantes and Giovanni Depau.

2 General description of Romance complementizers

2.1 The semantic relation of complementation and its syntactic realizations in Romance

Before giving an overview of Romance complementizers, we briefly introduce the semantic notion of complementation. We define *complementation* as the semantic situation that arises when a predicate takes a predication as one of its semantic arguments.

Fig. 1 illustrates the semantic relation linking the predicate 'remember' and the predication [*you arrive*] in a complementation situation:

(1) The semantic relation of complementation

I	*remember*	[*you arrive*]
ARG(entity)	PRED	ARG(predication)

This semantic situation has different syntactic realizations in Romance languages: sentential complementation (2); integration via an infinitive complement (3); integration via a WH-element which functions as a shared complement (4); integration via nominalization (5); integration with complementizer omission (6); and integration by juxtaposition (7). Let us examine them in detail.

Sentential complementation is the syntactic situation which arises when the argument is realized as a sentence, i.e. a finite clause, as in (2):

(2) Sentential complementation
French

Je me rappelle	*que*	[*tu*	*es*	*arrivé*]
I me remind.IND.PRS.1SG	COMP	you	be.IND.PRS.2SG	arrive.PST.PTCP

'I remember that you arrived'

When the argument predication is realized as an infinitive clause, we speak of *integration via infinitive complements*:

(3) Integration via infinitive complements
French

Je	*prévois*	*de*	[*partir*	*demain*]
I	plan.IND.PRS.1SG	COMP	leave.INF	tomorrow

'I plan to leave tomorrow'

A complementation phenomenon, which is quite frequent in Romance languages, is the integration via a wh-element which functions as a complement shared by both the predicate and the argument predication. Let us take, for example, the French sentence in (4), *Je me rappelle quand tu es arrivé* 'I remember when you arrived': as shown in our graphic representation of the dependency relations, the wh-element *quand* 'when' can be analysed as realizing both the head of the object of the verb *se rappeler* 'to remind oneself, to remember' and an adjunct of the verb *arriver* 'to arrive'. In such cases, we speak of *integration via a shared complement*:

(4) Integration via a shared complement
French

Je	*me rappelle*	[*quand*	*tu*	*es*	*arrivé*]
I	me remind.IND.PRS.1SG	when	you	be.PRS.2SG	arrive.PST.PTCP

'I remember when you arrived'

The argument predication can also be realized as a noun rather than as a verb. It is the case in instances of integration via nominalization (5):

(5) Integration via nominalization
French

Je	*me rappelle*	[*ton*	*arrivée*]
I	me remind.IND.PRS.1SG	your	arrival

'I remember your arrival'

Finally, complementation can be realized through two quite similar albeit non-coincident mechanisms of syntactic integration, which make no use of complementation markers: omission of the complementizer (6) and juxtaposition (7). An example of complementizer omission is given in (6), where the main predicate *credo* 'I believe' syntactically governs the verb *arrivare* 'to arrive' by imposing a subjunctive inflection: *sia arrivato* [be.SBJ.3SG arrived]. The two sentences are

linked by a dependency relation, in spite of the fact that there is no complementizer linking them (see below for a more detailed analysis).

(6) Integration with complementizer omission
Italian

Credo	[*sia*	*arrivato*]
believe.IND.PRS.1SG	be.SBJ.PRS.3SG	arrive.PST.PTCP

'I believe he has arrived'

The mechanism of juxtaposition is illustrated in (7): the argument predication *tu es arrivé* 'you arrived' is simply juxtaposed to the main predicate *je me rappelle* 'I remember', without any dependency marker linking the two parts of the sentence.

(7) Integration via juxtaposition
French

Je	*me rappelle*	[*tu*	*es*	*arrivé*]
I	me remind.IND.PRS.1SG	you	be.PRS.2SG	arrive.PST.PTCP

'I remember, you arrived'

The syntactic difference between integration with complementizer omission and integration via juxtaposition can be shown by applying a test of mobility. Such a test shows that while the postposition of the main predicate is not possible for the former (8), it is entirely possible for the latter (9):

Italian

(8) **Sia* *arrivato,* *credo*
be.SBJ.PRS.3SG arrive.PST.PTCP believe.PRS.1SG
'He has arrived, I believe'

French

(9) *Tu* *es* *arrivé,* *je* *me* *rappelle*
you be.IND.PRS.2SG arrive.PST.PTCP I me remind.IND.PRS.1SG
'you arrived, I remember'

It should be noted, incidentally, that the constructions illustrated so far show different degrees of *syntactic* dependency between the main verb and the com-

plement. The structures (2) through (6) are clear examples of microsyntactic dependency. By microsyntactic dependency, we mean the syntactic situation that arises when the existence, the category or the position of a word is determined by another word (see for example Tesnière 1959; Mel'cuk 1988; Kahane 2001).

In the structures (2) through (4) the very existence of the complementation markers *que* (in [2]), *de* (in [3]), *quand* (in [4]) is determined by the verbs of the main clause. In (5) the position of the object noun *ton arrivée* 'your arrival' is also determined by the main verb. In (6) the sentential complement is governed by the verb *credo* 'I believe', which imposes the subjunctive inflection on the verb *essere* in *sia arrivato* 'that he has arrived (lit. *he be arrived*)'. The syntactic dependency of a complement on a verb can be tested with the help of simple diagnostic tests, such as the equivalence of the dependent complement with a clitic pronoun (10):

(10) a. *Je me rappelle* 'I remember...
que tu es parti ...that you left'
quand tu es parti ...when you left'
ton départ ...your departure'
≡ *je me **le** rappelle* 'I remember **that'**

b. *Je prévois* 'I plan
de partir to leave'
≡ *je **le** prévois* 'I plan **that'**

c. *Credo sia arrivato* 'I believe he has left'
≡ ***lo** credo* 'I believe **that'**

or the fact that the dependent complement falls under the scope of the negation of the main verb (11):

(11) a. *Je ne me rappelle pas quand tu es arrivé/ que tu es arrivé/ ton arrivée*
'I don't remember when you arrived/ that you arrived/ your arrival'

b. *Je ne prévois pas de partir demain*
'I don't plan to leave tomorrow'

c. *Non credo sia arrivato*
'I don't think he has arrived'

The syntactic integration of the structure in (7) is slightly more problematic and it has been a hotly debated issue in the literature (Ross 1973; Schelfhout, Coppen & Oostdijk 2004; Dehé & Kavalova 2006; Ifantidou-Trouki 1993; Espinal 1991; Haegeman 1988). Following Kahane & Pietrandrea (2012), we claim that the integration between the main verb *je me rappelle* 'I remember' and the complement *tu es arrivé* 'you arrived' is not merely a matter of semantics or discourse, but a truly syntactic cohesion phenomenon. Kahane & Pietrandrea (2012) have proposed that, in spite of the fact that it is not possible to speak of a proper microsyntactic dependency (since the main verb does not control the existence, the position or

the form of the complement clause), the main verb is affected by some syntactic constraints – such as condition of adjacency or non-modifiability – which leads them to claim that a syntactic integration operates at a macrosyntactic level.

All in all, a number of syntactic realizations of the complementation situation are possible in Romance and complement markers and complementizers are not the only elements marking complementation in Romance. But let us further refine the notion of complementizer adopted in this article and thus specify the scope of our analysis.

2.2 Canonical and non-canonical complementizers

Following Noonan ([1985] 2007: 55), we define a complementizer as "a word, particle, clitic or affix, one of whose functions it is to identify [a clause] as a complement".

Two refinements to this notion are necessary, however.

Firstly, following the convention adopted in this volume, we define *canonical complementizers* as expressions that both conform to Noonan's definition and occur only in finite complements. Markers of integration such as the preposition *de* in (3), whose function is to mark the integration of infinitive complements, are therefore not taken into account in our paper.

Secondly, let us further examine Noonan's definition: Noonan states that complementizers are expressions *one of whose functions* is to identify a clause as a complement. Expressions functioning as complementizers can indeed have other functions in other contexts (as relativizers, adverbializers, etc. – see below, Section 2.3.). This taken into account, we restrict further our definition of complementizers as follows:

Canonical complementizers are expressions displaying the two following properties:

1. they have in other contexts among other possible functions, the function of identifying a clause as a complement.
2. their *only syntactic function in the sentence* is to identify a finite clause as a complement.

The second part of this definition leads us to exclude shared wh-complements such as 'when', 'where', 'how', 'what' and 'why' from the repertoire of canonical Romance complementizers. As we have seen in (4), these elements fulfill simultaneously two syntactic functions in one and the same sentence: they are used as complementation markers and they constitute a complement of the subordinate clause. They are thus *non-canonical complementizers*: they are *complementizers*,

because they are indeed used as complementation markers; they are *non-canonical*, because they fulfill two rather than one syntactic function in the sentence. Like canonical complementizers, they also have non-complementizer uses, as we will see in Section 2.3. These non-canonical complementizers make up a much larger category, with slightly more formal and functional variation between languages. Table 2 below lists a few non-canonical complementizers in Romance.

Table 2: 'How', 'what', 'when', 'where' and 'why' (non-canonical) complementizers in a few Romance languages

Complementizer form:	how	what	when	where	why
French	*comment*	*(ce) que / quoi*	*quand*	*où*	*pourquoi*
Italian	*come*	*che, che cosa, cosa*	*quando*	*dove*	*perché*
Portuguese	*como*	*o que*	*quando*	*onde*	*por quê*
Romanian	*cum*	*ce*	*cînd*	*unde*	*de ce*
Sardinian[4]	*comenti*	*ita*	*candu*	*du, aundi, inui…*	*poita*
Spanish	*cómo*	*qué*	*cuando*	*donde*	*por qué*

These non-canonical complementizers globally have similar distributions in each language, and occur in similar contexts in different languages, as illustrated by (12) below.

(12) *je (ne) sais (pas)* **quand/comment/pourquoi** *il est arrivé* (French)
(non) so **quando/come/perché** *è arrivato* (Italian)
(no) sé **cuando/cómo/por qué** *llegó* (Spanish)
(naõ) sei **quando/como/por quê** *chegou* (Portuguese)
(nu) ştiu **cînd/cum/de ce** *a sosit* (Romanian)
(no) iscu **candu/comenti/poita** *est arribau* (Sardinian)
'I (don't) know when/how/why he arrived'

The question is open as to whether relative pronouns, like the equivalents of *what* in (13) below, should be included in this category. Indeed, relative pronouns share some properties with non-canonical complementizers: they both identify the clause as a complement and have a syntactic function in the subordinate clause. Unlike non-canonical complementizers, though, relativizers often have an anaphoric relation to their so-called "antecedent" in the main clause, except in case of "headless relativizers" as in free relatives like (13) below. This is the

4 For all languages, we refer as much as possible to the most widely accepted 'standard' variant. This is of course more problematic for some languages than others; very much so for Sardinian (see e.g. Mensching 2004:14). The Sardinian data presented here do not render justice to the enormous variation between Sardinian dialects, but that is not our goal.

most common argument for a distinction between relativizers and non-canonical complementizers, which is, besides, not unanimously subscribed to by all scholars (see for example Kayne 2005: 13 for the Generative approach, but also Le Goffic 1993: 525 in another syntactic framework).

(13) *je sais* **ce que** *je veux* (French)
je sais **quoi** *je veux* (Substandard French)
so **(che) cosa** *voglio* (Italian)
so **che** *voglio* (Colloquial Italian)
sé **qué** *quiero* (Spanish)
sei **o que** *quero* (Portuguese)
ştiu **ce** *vreu* (Romanian)
iscu **ita** *bolu* (Sardinian)
'I know what I want'

In almost all Romance languages, canonical complementizers are restricted to two series, one functionally close to the 'that' complementizer in English, the other to the 'if' complementizer (Table 3 below).

Table 3: 'that' and 'if' complementizers in Romance languages (except Romanian)

	'that' complementizer	**'if' complementizer**
Spanish	*que*	*si*
Portuguese	*que*	*se*
Sardinian	*chi*	*si*
Italian	*che*	*se*
French	*que*	*si*

Typical examples of the 'that' complementizer are given in (14) below, and in (15) for the 'if' complementizer.

(14) *je sais* **qu**'*il est arrivé* (French)
so **che** *è arrivato* (Italian)
sé **que** *llegó* (Spanish)
sei **que** *chegou* (Portuguese)
ştiu **că** *a sosit* (Romanian)
iscu **chi** *est arribau* (Sardinian)
'I know that he has arrived'

(15) *je me demande* **s'***il est arrivé* (French)
mi chiedo **se** *è arrivato* (Italian)
me pregunto **si** *llegó* (Spanish)
pergunto-me **se** *chegou* (Portuguese)
mă întreb **dacă** *a sosit* (Romanian)
mi preguntu **si** *est arribau* (Sardinian)
'I wonder whether he has arrived'

Note that Romanian presents a different configuration, with three canonical complementizers, two 'that' complementizers and one 'if' complementizer (Table 4 below): only Romanian has a different complementizer in (14) and (16).

Table 4: 'that' and 'if' complementizers in Romanian

	'that'$_1$-complementizer	'that'$_2$-complementizer	'if' complementizer
Romanian	*că*	*să*	*dacă*

(16) *je veux* **qu'** *il vienne* (French)
voglio **che** *venga* (Italian)
quiero **que** *venga* (Spanish)
quero **que** *venha* (Portuguese)
vreu **să** *vină* (Romanian)
bolu **chi** *bèngiat* (Sardinian)
(I) want **that** (he) come.SBJV.PRS3SG
'I want him to come'

2.3 Non-complementizer use of canonical complementizers

It should be noted that canonical complementizers can have other uses in Romance besides complementation.

'That' forms have a relativizer function in all Romance languages (except Romanian), as illustrated by the Portuguese example below:

Portuguese

(17) *a cura de austeridade,* ***que*** *(...) irá cortar nos subsídios de férias e de Natal*
'The austerity measures, **which** will cut into the vacation and Christmas bonuses'

They can also be used to introduce adverbial clauses, for example purposive clauses (18), (19):

Italian

(18) *vieni qua **che** ti spiego*
'Come here and I'll explain'

Portuguese

(19) *anda cá, **que** quero falar contigo*
'Come here, I want to talk to you'

Another function of complementizer forms is to introduce non-assertive de-subordinated main sentences (Frajzyngier 1995), as in the French, Romanian and Italian examples below. These sentences can be of various types, jussive, interrogative (apparently only in Italian) and optative:

jussive, French

(20) ***Qu'**il vienne!*
'Let him come!'

jussive, Romanian

(21) ***Să** vină!*
'Let him come!'

interrogative, Italian

(22) ***Che** hai sonno?*
'Are you asleep?'

optative, French

(23) ***Si** seulement je pouvais lui manquer*
'If only he could miss me'

optative, Romanian

(24) ***Dacă** (măcar) ar veni*
'If only he would come'

A *quotative* function of 'that' forms is found apparently only in French (see (25) below):

French

(25) *Cette femme est incroyable! "Allez, Allez!"* ***qu'****elle dit à la journaliste*
'This woman is incredible. "Go ahead" she says to the journalist'. (Tweet)

'If' forms are used in hypothetical constructions, more precisely, they introduce conditional protasis clauses. This adverbializer use, illustrated here for Spanish (26), is found in all Romance languages:

Spanish

(26) ***Si*** *tú no vuelves, se secarán todos los mares*
'If you don't come back, all the seas will dry up'

2.4 Combinability issues

The combination of complementizers with other subordination markers is a marginal phenomenon in Romance languages. It appears mostly in regional and/or substandard, informal variants, in which the 'that' complementizers can be combined with various complementation markers, such as 'when' markers. Not all combinations of non-canonical complementizer + 'that' complementizer are attested, and most of them are clearly substandard and very rare. However, we found in our corpora constructions such as 'when that' (French *quand que*, Italian *quando che*, Portuguese *quando que*, Spanish *cuando que*, Sardinian *candu chi*) or 'how that' (French *comme que*, Portuguese and Spanish *como que*, Romanian *cum că*, Sardinian *comenti chi*), albeit often with very low frequencies. Examples below illustrate the constructions 'when' + 'that' for French (27) and Italian (28), and 'how' + 'that' for Romanian (29) and Sardinian (30):

Louisiana French

(27) *Je le ferai* ***quand*** ***que*** *j'irai à la ville.*
I CL do.IND.FUT.1SG **when** **that** I.go.IND.FUT.1SG to the town
'I'll do it when I'll go into town'
(Ditchy 1997: 173)

Regional Italian (Rome)

(28) ***Quanno*** ***che*** *'ncomincia* *la* *partita* *ogni* *tifosetta* *se*
when **that** begin.IND.PRS.3SG the.F match every fan.DIM.F.SG REFL
fà *ardita*
make.IND.PRS.3SG bold.F.SG
'When the match begins, every little fan becomes bolder'

Romanian

(29) *Mi-a* *spus* ***cum*** ***că*** *ar*
me.DAT-have.AUX.3SG tell.PST.PTCP **like** **that** have.AUX.COND.3SG
fi *plecat*
be.AUX.INF leave.PST.PTCP
'(s)he told me that (s)he would be gone'

Sardinian

(30) *Bivi* ***comenti*** ***chi*** *no* *depis* *mòrriri*
live.IMP.2SG **like** **that** NEG must.IND.PRS.2SG die.INF
'Live as if you were to die tomorrow'

It should be mentioned that the semantic change induced by the adjunction of a 'that' complementizer in these constructions is not necessarily the same in different instances, including clear semantic contrasts such as the use of *when + that* to introduce conditionals (in the Emilian dialect of Italian, cf. Mazzoleni 1997: 202): these changes are thus not easily identified. Such constructions seem to suggest that, in some cases, the 'that' complementizer has grammaticalized to the point of being semantically neutral, and a obligatory marker of subordination (see Section 8), while in others it has retained a semantic value.

3 General semantic functions of Romance complementizers

Having isolated the class of canonical complementizers and studied their syntactic properties, we focus in the next sections on their functional aspects: we first observe the general illocutionary (Section 4.1.) and semantic properties (Section 4.2.) of complementizers, and then analyse the semantic difference between 'if' and 'that' complementizers (Section 5).

3.1 Complementizers as markers of illocutionary dependency

The first question to be answered is: what is the semantic contribution of a complementizer to the sentence? In other words, we want to determine to what extent the presence vs. absence of a complementizer (irrespective of its nature) may change the meaning of a sentence.

We claim that *complementizers have the general function of marking the argument predication as illocutionarily dependent on the main predicate*. In order to further specify this notion, we refer to Cristofaro's (2005) functional definition of subordination. According to her, all cases of subordination presuppose an "asymmetry assumption": "a cognitive asymmetry is established between linked SoAs[5], such that the profile of one of the two (henceforth, the main SoA) overrides that of the other (henceforth, the dependent SoA). This is equivalent to saying that the dependent SoA is (pragmatically) non-asserted, while the main one is (pragmatically) asserted" (Cristofaro 2005: 35). Since complementation is a particular case of subordination, following Cristofaro we may say that, in a complementation situation, the argument predication is non-asserted, whereas the situation depicted by the main predicate is asserted.

This asymmetry has been diversely addressed in the literature (see for example Cristofaro 2005; Boye & Harder 2009; Kahane & Pietrandrea 2012) and a battery of tests has been elaborated to distinguish between illocutionarily independent and illocutionarily dependent situations representations. The test in (31), for example, shows a canonical semantic situation of complementation, in which the main predicate bears the illocutionary force of assertiveness, whereas the argument predication lacks it: in the discourse negotiation (B, in example 31) the interlocutor challenges the truth of the assertion made by A, and it is clear that his challenge focuses on the main predicate, *je t'ai dit* 'I told you', rather than on the argument predication *il est venu* 'he came'.

French

(31) A: *Je t'ai dit qu'il est venu*
'I told you that he has come'
B: *Ce n'est pas vrai ! (que tu me l'as dit)*
'It's not true (that you told me)'
? *Ce n'est pas vrai ! (qu'il est venu)*
'It's not true (that he came)'

5 SoA(s): state-of-affairs.

Similarly, a tag question such as *n'est-ce-pas ?* 'isn't it?', whose function is to modify the illocutionary force of the anchor sentence, has scope on the main predicate – i.e. the part of the sentence that has an illocutionary force which can be modified – rather than on the argument predication, which lacks illocutionary force – see (32):

French

(32) *Je t'ai dit qu'il est venu, n'est-ce-pas?*
'I told you that he has come, didn't I?'
'*I told you that he has come, didn't he?'

Concerning Romance complementation via juxtaposition, we claim that the absence of the complementizer changes the direction of the illocutionary dependency between main predicate and argument predication. Let us examine a sentence such as (33). In (33), there is no syntactic integration between the two parts of the sentence. However, a semantic integration between the two parts can be detected: the predicate *je t'ai dit* 'I told you' takes in fact as semantic object the predication *il est venu* 'he came'. From a semantic point of view, *il est venu* is thus dependent on *je t'ai dit*. But from an illocutionary point of view, we cannot say that the semantic object *il est venu* is dependent on the main predicate *je t'ai dit*. Rather, as the test shows, it is the semantic object that naturally bears the illocutionary force of the utterance:

French

(33) A: *Il est venu, je t'ai dit*
'He has come, I told you'
B: *Ce n'est pas vrai ! (qu'il est venu)*
'It's not true (that he has come)'
B: ?* *Ce n'est pas vrai ! (que tu me l'as dit)*
'It's not true (that you told me)'

A similar test can be applied to complementizer omission in Italian. Let us compare for example (34) and (35):

Italian

(34) *Penso* *che* *è* *simpatico*
think.IND.PRS.1SG that be.IND.PRS.3SG funny
'I think that he is funny'

(35) *Penso* *sia* *simpatico*
think.IND.PRS.1SG be.SBJV.PRS.3SG funny
'I think he is funny'

In (34), the use of a so-called *recteur faible* ('weak head', cf. Blanche-Benveniste 1989; Blanche-Benveniste & Willems 2007), *penso* 'I think', followed by the complementizer *che* 'that', leaves open the ambiguity concerning the direction of the illocutionary dependency, as is shown by the test in (34'):

Italian

(34') A: *Penso che è simpatico*
'I think that he is funny'
B: *non mi sembra (che lo pensi)*
'It does not seem to me (that you think this)'
B: *non mi sembra (che è simpatico)*
'It does not seem to me (that he is funny)'

In (35), instead, the use of the same *recteur faible*, *penso* 'I think', without any complementizer imposes the reading according to which it is the semantically dependent part of the sentence (the semantic object *è simpatico* 'he is nice') that bears the illocutionary force of the utterance, as is shown in (36):

Italian

(36) A: *Penso sia simpatico*
'I think he is funny'
B: **non mi sembra (che lo pensi)*
'It does not seem to me (that you think that)'
B: *non mi sembra (che è simpatico)*
'It does not seem to me (that he is nice)'

In sum, we claim that one of the main general functions of complementizers in Romance is to mark the argument predication as not only syntactically and semantically, but also illocutionarily dependent on the main predicate.

3.2 Complementizers as markers of modality

Along with the general function of markers of illocutionary dependency, it is well known that at least some complementizers may have modal functions. According to Frajzyngier (1995), the focal function of complementizers is actually a modal one (which has been diachronically eroded in many languages). Japanese, for

example, has three complementizers: *no, koto* and *to*. In many cases the choice among the three forms is determined by the verb of the matrix clause. In contexts in which the choice is open, *to* indicates that the speaker does not believe in the truth of the proposition represented in the complement. By contrast, neither *no* nor *koto* has this modal value:

Japanese

(37) *John wa Mary ga sin-da to sinzi-na-katta*
John TOP Mary SBJV die-PST COMP believe-NEG-PST
'John did not believe that Mary was dead' (she might or might not have been)
(Kuno 1973: 216 – quoted in Palmer 1986: 148)

(38) *John wa Mary ga sin-da koto o sinzi-na-katta*
John TOP Mary SBJV die-PST COMP OBJ believe-NEG-PST
'John did not believe that Mary was dead' (she was)
(Kuno 1973: 216– quoted in Palmer 1986: 148)

In Jacaltec (Mayan), different complementizers can mark different degrees of certainty. As noted by Craig (1977: 267), "the use of [the complementizer] *chubil* denotes a high degree of credibility or certainty, and the use of [the complementizer] *tato* introduces a notion of disbelief or reservation about a hearsay". She provides the (much quoted, see e.g. Noonan 2007: 48) examples below to illustrate this opposition, and remarks that in (39) "the source of information (...) is taken to be an unreliable source of information", while in (40) "the source of information is felt to be a reliable one" (Craig 1977: 268):

Jacaltec

(39) *Xal naj tato chuluj naj presidente*
said DET COMP will.come DET president
'He said that the president is going to come' [unreliable source]
(Craig 1977: 268)

(40) *Xal naj alcal chubil chuluj naj presidente*
said DET alcalde COMP will.come DET president
'The alcalde said that the president would come' [reliable source]
(Craig 1977: 268)

In Kinya Rwanda, a Bantu language, there are three complementizers: *ko, ngo* and *kongo*. *Ko* indicates that the speaker has a neutral attitude with respect to the truth of the complement, while *ngo* and *kongo* indicate that the speaker doubts the truth of the complement (on the basis of some direct and indirect evidential justification, respectively) (Givón 1982: 30):

Kinya Rwanda

(41) *yavuze* *ko* *azaaza*
he.PST.say that he.FUT.come
'He said that he'd come'
(Givón 1982: 30)

(42) *yavuzu* *ngo* *azaaza*
he.PST.say that he.FUT.come
'He said that he'd come' (but I doubt it)
(Givón 1982: 30)

(43) *yavuze* *kongo* *azaaza*
he.PST.say that he.FUT.come
'He said that he'd come' (but I have indirect evidence which makes me doubt it)
(Givón 1982: 30)

In addition, Kinya Rwanda has a complementizer, *-niiba-*, marking the speaker's genuine (i.e. not evidentially supported) uncertainty about the truth of the complement clause (Givón 1982: 30).

As for European languages, it is well known that in some Slavic languages as well, some complementizers can have an epistemic meaning. Jakobson (1957), for example, reports their presence in Russian, where for instance *budto* 'as if' can be used as a complementizer and mark the speaker's doubts concerning the truthfulness of a statement. Similarly, Frajzyngier shows that the Polish construction *niby* – formed by the epistemic marker *ni* and the irrealis enclitic *by,* always occurring in clause-initial position, i.e. in the typical position of Polish complementizers – expresses the speaker's doubt about the truth of the proposition:

Polish

(44) *Ni-by* *był* *w* *Warszawie*
COMP be.3M.SG.PFV in Warsaw.DAT
'Apparently he was in Warsaw'
(Frajzyngier 1995: 495)

This complementizer can co-occur with the *de dicto* complementizer *że* and introduce complements concerning situations presented as uncertain by the speaker:

(45) *Powiedział* *że* *ni-by* *był* *w* *Warszawie*
say.PFV.PST.3SG.M COMP COMP be.3SG.M.PFV in Warsaw.DAT
'He said that he was in Warsaw' (I have my doubts about it)
(Frajzyngier 1995: 496)

Nordström (2010) has shown that, in Germanic languages, 'that' complements can depict either logically realis or logically irrealis situations, i.e., either situations in an "actualized or certain fact of reality" (Elliott 2000: 66–67), or situations portrayed "as purely within the realm of thought, knowable only through imagination" (Mithun 1995: 173), whereas 'if' complements are always irrealis.

Nordström & Boye (this volume) refine this proposal, by characterizing the contrast between 'that' and 'if' complementizers in Germanic languages as an epistemic modal contrast: 'that' complements are epistemically neutral, whereas 'if' complements encode the "uncertainty about the proposition expressed by the complement". In the next section, we closely examine the difference between 'that' and 'if' complements in Romance languages. We argue that the semantic contrast between the two complementizers is better described as a contrast in factuality than as a modal contrast.

We first focus our analysis on French, Italian, Spanish, Portuguese and Sardinian, and go on to examine closely the case of Romanian, which has a system of three complementizers quite different from the other Romance languages.

4 Semantics of 'that' and 'if' complements in Romance languages

4.1 Distribution: Predicates taking either 'that' or 'if' complements

The semantic difference between 'if' and 'that' is visible in their distribution, especially in the semantic contrast between predicates taking 'that' complements and predicates taking 'if' complements.

In order to examine this contrast, we searched, through the Sketch Engine query system (http://www.sketchengine.co.uk; Kilgariff et al. 2004), five large (written-language) corpora: the TenTen Corpora for French (frTenTen 10,666,617,369 words), Italian (ItTenTen 2,588,873,046 words), Spanish (esTenTen 2,103,770,763 words) and Portuguese (ptTenTen 2,757,635,105 words) and the Romanian Web as a Corpus for Romanian (RoWaC, 44,729,032 words). These corpora are created by Web crawling, processed with cleaning and de-duplication tools and lemmatized; the French, Spanish and Italian corpora are also tagged for part of speech. Although these corpora cannot be regarded as representative of the languages in focus, their important size allowed us to identify and measure the salience of some very general patterns. This was particularly interesting for

our purpose, i.e., identifying and comparing general tendencies in Romance languages.

In these corpora we searched for the most frequent verbs immediately preceding 'that' and 'if' complementizers, in all five languages. We could not search for all verbs governing 'that' and 'if' complementizers, as dependency relations are not encoded in our corpora and we could only rely on a positional criterion. Nevertheless, we verified and eliminated all the false positives (i.e., the verbs immediately preceding 'that' and 'if' complementizers that do not syntactically govern them) and the results we obtained are quite clear and consistent for the different languages under examination.

Let us consider in detail the following tables, which show the 12 most frequent verbs immediately preceding and governing 'that' and 'if' complementizers in Italian, French, Spanish and Portuguese. We study the distribution of Romanian, which is quite different from the others, in Section 6.1.

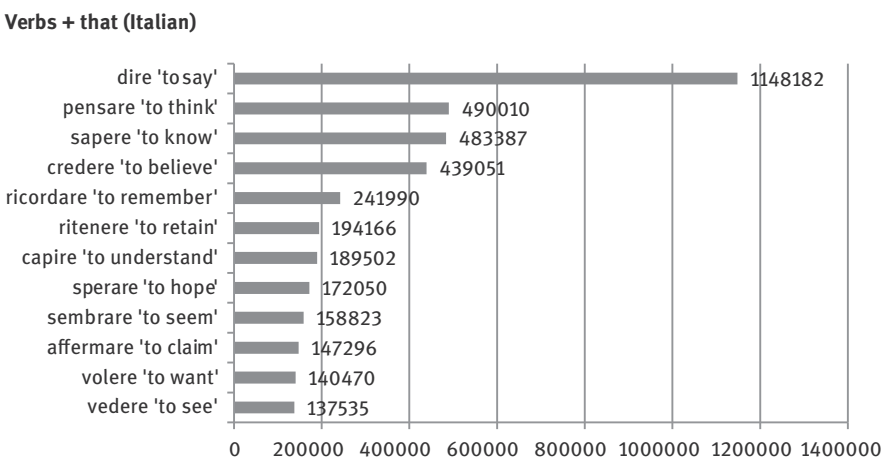


Figure 1: Most frequent verbs introducing complement clauses with 'that' in Italian

Verbs + if (Italian)

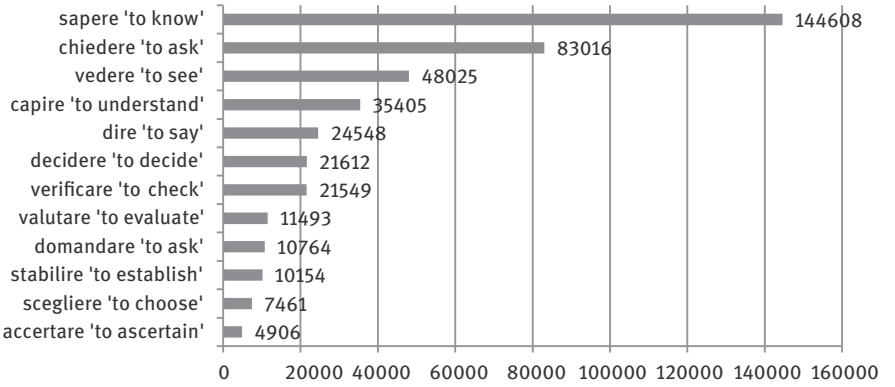


Figure 2: Most frequent verbs introducing complement clauses with 'if' in Italian

Verbs + that (French)

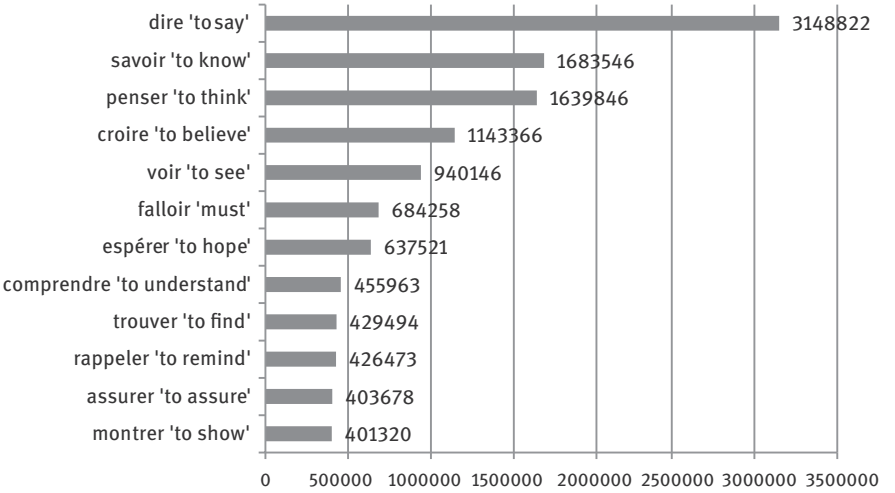


Figure 3: Most frequent verbs introducing complement clauses with 'that' in French

Verbs + if (French)

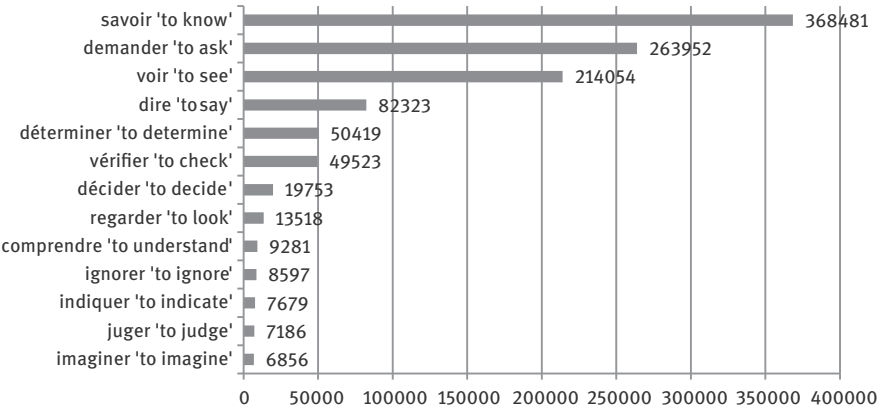


Figure 4: Most frequent verbs introducing complement clauses with 'if' in French

Verbs + that (Spanish)

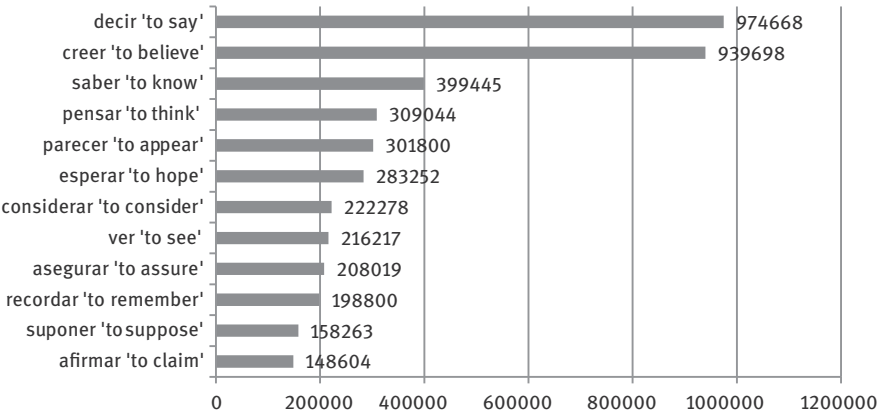


Figure 5: Most frequent verbs introducing complement clauses with 'that' in Spanish

Verbs + if (Spanish)

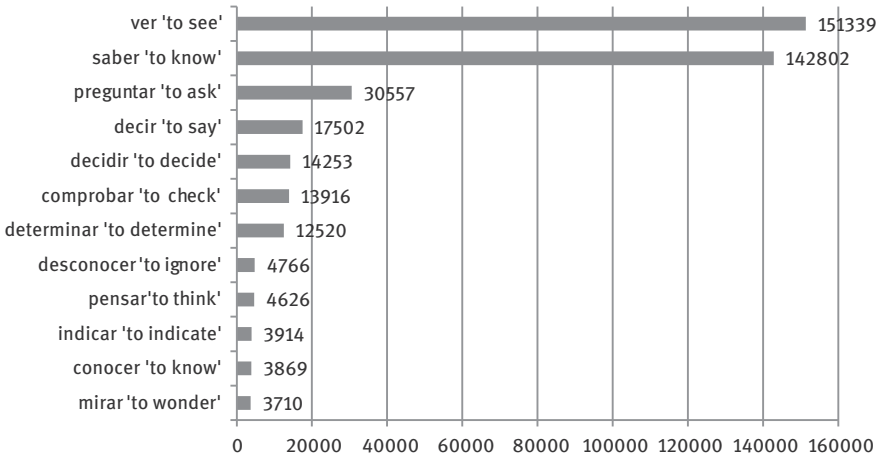


Figure 6: Most frequent verbs introducing complement clauses with 'if' in Spanish

Verbs + that (Portuguese)

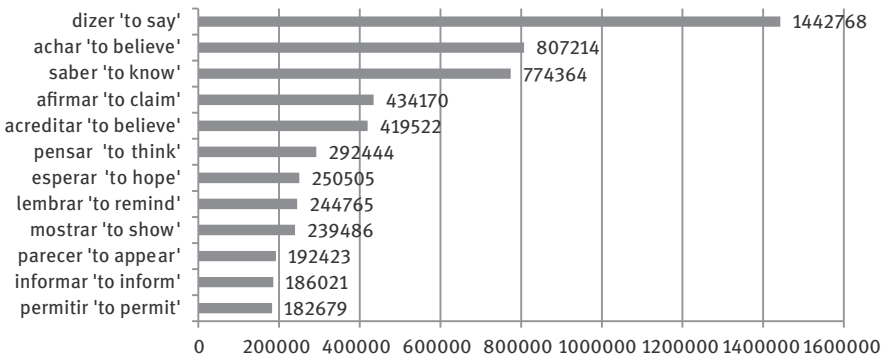


Figure 7: Most frequent verbs introducing complement clauses with 'that' in Portuguese

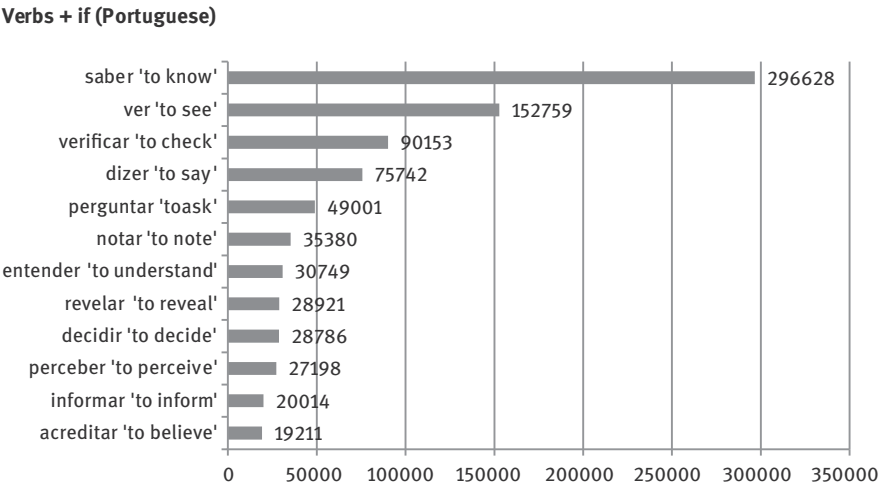


Figure 8: Most frequent verbs introducing complement clauses with 'if' in Portuguese

The first fact immediately observable is that, in Romance languages, there are different series of predicates. One series takes only 'that' complements, and another one takes only 'if' complements.

'Say', 'think', 'know', 'believe' are the four most frequent verbs governing 'that' complements in Italian, French, Spanish and Portuguese (the situation of Romanian is slightly different, see Section 6.1.). Whereas the verbs 'ask', 'see', 'know' are the most frequent verbs governing 'if' complements.

At first sight, it may seem that some predicates such as the verbs 'know', 'say', 'understand', 'see' govern, albeit unevenly, both 'if' and 'that' predicates. A closer look at the data shows, however, that these highly polysemic predicates are employed in different contexts or with different meanings depending on whether they govern 'that' or 'if' complementizers. For example, the verb 'know' only appears in negative (46) or non-factual contexts (47) when governing 'if' complements, whereas it can occur in factual contexts (48) when governing 'that' complements:

Spanish

(46) ***no sé si*** *puedo viajar con un visado de turista*
'I don't know if I can travel with a tourist visa'

Spanish

(47) ***queria saber si*** *puedo viajar con un visado de turista*
'I wanted to know if I can travel with a tourist visa'

Spanish

(48) ***yo sé que*** *puedo viajar con un visado de turista*
'I know that I can travel with a tourist visa'

Similarly, the predicate 'say' occurs in negative or non-factual contexts (49–50) when governing 'if' complements and in factual contexts when governing 'that' complements (51):

Italian

(49) ***Non dice se*** *esiste una cura in tal senso*
'It (/he) doesn't say if there exists a cure in that sense'

Italian

(50) ***Potrebbe gentilmente dirmi se*** *esiste una cura in tal senso*
'Could you kindly tell me if there exists a cure in that sense'

Italian

(51) ***Dice che*** *esiste una cura in tal senso*
'He says that there exists a cure in that sense'

The verbs 'understand' and 'see' are cognitive verbs which can be classified, as for Aktionsart (Vendler 1967), as accomplishment verbs, i.e., as verbs whose internal temporal nature can be analysed as being durative and telic. Interestingly, when these verbs govern 'if' complements, it is the durative phase that is semantically selected (52), (54), whereas when they govern 'that' complements it is the telicity that it is selected (53), (55):

French

(52) ***il est en train de voir si*** *cela est possible et pas dangereux*
'he is trying to see if this is possible and not dangerous'

French

(53) ***j' ai vu que*** *c'était possible et pas dangereux*
'I have seen that this is possible and not dangerous'

French

(54) ***j' essaie de comprendre si*** *vous êtes un souverainiste comme on dit*
'I am trying to see if you are a partisan of sovereignty, as people say'

French

(55) ***j' ai compris que*** *vous êtes un souverainiste comme on dit*
'I have understood that you are a partisan of sovereignty, as people say'

In light of these results, we may propose a first rough generalization concerning the semantics of the predicates governing 'if' and 'that' complementizers.

Predicates governing 'if' complementizers can be regarded as a specific type of cognitive predicates, i.e. uncertainty predicates. An uncertainty predicate presents its complement as *not known* by the speaker (or the subject) of the utterance: either the speaker questions its truth value, or he claims not to know it, or he is depicted as being in the middle of a search for truth (like, for example, when a cognitive verb in its durative phase is used). The complements of these predicates correspond to what traditional grammar calls "embedded questions": these complements are always presented as non-factual by the speaker.

Predicates governing 'that' complementizers belong, instead, to the larger class, already identified by Givón ([1984] 2001:154), of PCU predicates: perception ('see'), cognition ('think', 'know') and utterance predicates ('say', 'tell'). Verbs belonging to the class of perception, cognition and utterance predicates may or may not cast non-factuality on the embedded clause (e.g., the verb 'know' entails the factuality of its complements, whereas the verb 'think' does not)[6].

Concerning the modal value of 'that' and 'if' complementizers, we may therefore conclude that in Romance languages 'if' complements are always non-factual, whereas 'that' complements are unmarked for factuality.

In the following sections, we refine this general statement by showing

1. which type of non-factuality is encoded by 'if' complements;
2. the role played by mood in determining the factuality status of 'that' complements;

6 For the notion of factuality we refer to Eliott (2000), Mithun (1995).

3. the areal exceptions to the general Romance rule constituted by French on one side and Romanian on the other side.

4.2 The non-exclusion of factuality value of 'if' complements

The occurrence of 'if' complements within the scope of uncertainty predicates entails that 'if' complements have a specific type of non-factual meaning, i.e. what has been called a non-exclusion-of-factuality, or NEF meaning (Masini & Pietrandrea 2010; Pietrandrea 2011). Let us examine closely this conceptual notion, thoroughly described by Pietrandrea (2011).

As argued by Pietrandrea (2011: 186), non-factuality can be seen as a super-category encompassing several domains concerned with qualifying the actualization of a situation. A situation may be presented as not grounded in the perceivable reality for at least three different semantic reasons:

(a) The situation is qualified as not real, i.e. as not occurring or not having occurred in the actual world: a negated situation, for example. This sense of non-factuality may be labeled *counterfactuality*.
(b) The situation is introduced without specifying its spatio-temporal references in the actual world. The situation may be an a-temporal non-referential or a habitual situation: this sense may be labeled *non-referentiality* or, more generally, *non-factuality* (for the relation between non-referentiality and non-factuality, see Givón 2001: 303).
(c) The situation is presented as being just an eventuality, in other terms the factuality of the situation is not to be excluded among other options. This third sense of non-factuality has been labeled by Masini & Pietrandrea (2010) and Pietrandrea (2011) *non-exclusion of factuality*, or *NEF*.

Non-exclusion of factuality is a defining component of some typically non-factual situations: hypothetical conditionals, concessive conditionals, some optatives, some recommendations and alternative relations (Pietrandrea 2011). In all these situations, the non-factual situation is portrayed as being an eventuality, i.e., as just one of a set of mutually exclusive alternatives.

Hypothetical conditional

(56) *If it rains I will stay at home (and if it does not rain I will go out)*
(Prandi 2004; Dancygier &Sweetser 2005)

Concessive conditional

(57) *Even if we do not get any financial support, we will go ahead with our project*
(Haspelmath & König 1998)

Alternative relations

(58) *Either I walk or I go by bus*
(Mauri 2008)

From a communicative point of view, non-exclusion of factuality expressions are not modal in nature: they suspend the commitment of the speaker rather than qualifying it (cf. Pietrandrea 2011; see also Bellert 1977: 344–345, who introduced this important distinction which had been quite neglected in the literature). In this sense, non-exclusion of factuality expressions convey a type of non-factuality that is quite easily distinguishable from the non-factuality encoded by epistemic, deontic and volitional predicates. Epistemic, deontic, volitional predicates *qualify* the commitment of the speaker toward a non-factual situation (i.e., toward a non-referential situation), whereas non-exclusion of factuality meanings do not qualify the speaker commitment.

From a semantic and syntactic point of view, it is important to stress that non-exclusion of factuality meanings in general and 'if' complements in particular depict semantic situations implying alternatives. We have mentioned indeed that an eventuality can be defined as one of a set of mutually exclusive options. It is always possible for 'if' complements (as well as for any other non-exclusion of factuality expression, Pietrandrea 2011) to make explicit these other options, as shown by the following examples:

(59) *je me demande s'il est arrivé* ***ou non*** (French)
mi chiedo se è arrivato ***o no*** (Italian)
me pregunto si llegó ***o no*** (Spanish)
pergunto-me se chegou ***ou não*** (Portuguese)
mă întreb dacă a sosit ***sau nu*** (Romanian)
mi preguntu si est arribau ***o no*** (Sardinian)
'I wonder whether he has arrived **or not**'

(60) *je ne sais pas s'il est arrivé* ***ou non*** (French)
non so se è arrivato ***o no*** (Italian)
no sé si llegó ***o no*** (Spanish)
não sei se chegou ***ou não*** (Portuguese)
nu ştiu dacă a sosit ***sau nu*** (Romanian)
no ischo si est arribau ***o no*** (Sardinian)
'I don't know whether he has arrived **or not**'

(61) *je n'ai pas compris s'il est arrivé* ***ou non*** (French)
non ho capito se è arrivato ***o no*** (Italian)
no entendí si llegó ***o no*** (Spanish)
não percebi se chegou ***ou não*** (Portuguese)
n-am înțeles dacă a sosit ***sau nu*** (Romanian)
no appo cumpresu si est arribau ***o no*** (Sardinian)
'I didn't understand whether he has arrived **or not**'

The data from corpora show that the alternatives are indeed often made explicit within 'if' complements.

Italian

(62) *non ho capito* ***se rispondono tramite email oppure contattano telefonicamente***
'I didn't understand whether they answer by e-mail or (if) they contact you by telephone'

French

(63) *je ne sais pas* ***si tu parles à moi ou bien à skander09***
'I don't know if you are talking to me or to skander09'

Spanish

(64) *no he visto* ***si es optico o laser***
'I didn't see if it is optical or laser'

Portuguese

(65) *uma outra que não sei* ***se é cobra ou se é peixe***
'another one which I don't know if it is a cobra or a fish'

It should be said that the literature often regards 'if' complements as expressing the modal notion of doubt (Nordström 2010). The Romance data, however, lead us to question such an interpretation. As the following examples show, predicates of doubt in Romance generally cannot be followed by 'if' complements.

(66) **je doute s'il est arrivé* (French)
**dubito se è arrivato* (Italian)
dudo si llegó but also ^{ok}*dudo que llegase* (Spanish)
**dúvido se chegou / chegasse* (Portuguese)
**mă îndoiesc dacă a sosit* (Romanian)
**dubbido si est arribau* (Sardinian)
'I doubt if he has arrived'

Dubitative situation, indeed, do not imply alternatives:

(67) **je doute s'il est arrivé ou pas* (French)
**dubito se è arrivato o no* (Italian)
**dudo si llegó o no* (Spanish)
**dúvido se chegou ou não* (Portuguese)
**mă îndoiesc dacă a sosit sau nu* (Romanian)
**dubbido si est arribau o no* (Sardinian)
'*I doubt that he has come or not'

Dubitative situations are modal in nature and qualify the commitment of the speaker rather than suspending it.

In order to corroborate our hypothesis that 'if' complements serve the purpose of suspending the commitment of the speaker by putting forward the depicted situation as but one of a set of mutually exclusive situations, let us examine the possibility of modification of 'if' complements by non-exclusion of factuality and epistemic adverbs.

4.3 Constraints on the adverbial modification of 'if' complements

Let us start by making a semantic distinction between modal epistemic and non-exclusion of factuality adverbs. Semantically, we may define a modal epistemic adverb as an adverb which qualifies the commitment of the speaker toward what he says. For example, let us take the English epistemic adverb *probably*: by using this adverb (65), the speaker qualifies his/her commitment by saying that he/she puts forward the content of the utterance as probable.

(68) *This means that he is **probably** interested in her, which is why he hangs around so much.*

A non-exclusion of factuality adverb, instead, serves the purpose of suspending, rather than qualifying, the commitment of the speaker. For example, by using the English non-exclusion of factuality adverb *by chance* in (66), the speaker does not qualify his commitment, but simply presents the depicted situation as one option which is not to be excluded:

(69) *If **by chance** you have already visited Craigendarroch under the terms of this special invitation, I very much regret that we are unable to make the offer a second time (BNC corpus)*

Using a non-exclusion of factuality adverb, the speaker can present a situation as non-factual without having to qualify his commitment as to the likelihood of the situation.

In order to better appreciate the distinction between non-exclusion of factuality and epistemic adverbs, let us examine the compatibility of both kinds of adverbs with protases of hypothetical constructions.

As observed by Bellert (1977), the protasis of a conditional construction serves the purpose of suspending the commitment of the speaker, therefore, these constructions cannot host epistemic adverbs, since it is not possible for the speaker to suspend and qualify his/her commitment at the same time. On the contrary, NEF adverbs can occur in the protases of hypothetical constructions.

So for example, we have for Romance languages the following situation: the Italian adverb *magari* 'perhaps' (Pietrandrea 2010), the French adverb *des fois* 'perhaps', the Spanish adverb *a lo mejor* 'perhaps' (Cornillie 2010), the Portuguese adverb *por ventura* 'perhaps', the Romanian adverb *cumva* 'perhaps' can all occur in protases, whereas the modal epistemic adverbs corresponding to *probably*, i.e., *probabilmente* (Italian), *probablement* (French), *probablemente* (Spanish), *provavelmente* (Portuguese), *probabil* (Romanian) do not occur in similar contexts:

Italian

(70) *un cellulare gli potrebbe tornare utile, così,* ***{ nel caso in cui / se } magari*** *un giorno non riuscisse ad alzarsi dal letto per chissà quale motivo.*
'a cell phone might turn out useful to him, like, { in case / if } perhaps one day he didn't manage to get up for whatever reason'

(71) **un cellulare gli potrebbe tornare utile, così,* ***{ nel caso in cui / se } probabilmente*** *un giorno non riuscisse ad alzarsi dal letto per chissà quale motivo.*
'a cell phone might turn out useful to him, like, { in case / if } probably one day he didn't manage to get up for whatever reason'

French

(72) *Super, vous pouvez ramener ça* ***{ au cas où / si } des fois*** *mon disque implose ou une autre bêtise du genre*
'Great, you can take this back { in case / if } perhaps my disk implodes or some other crazy thing happens'

(73) **Super, vous pouvez ramener ça* ***{ au cas où/ si } probablement*** *mon disque implose ou une autre bêtise du genre*
'Great, you can take this back { in case / if } probably my disk implodes or some other crazy thing happens'

Spanish

(74) ***Si a lo mejor/ *probablemente*** *no tiene puesto ningún carrete, me doy por satisfecha.*
'If { perhaps / *probably } there's no film inside, I'll consider myself satisfied' (on-line dictionary example, cf. http://en.bab.la/dictionary/english-spanish/film-festival)

Portuguese

(75) *Recordo-me de uma aula. Nessa altura nós tínhamos falta e havia um contínuo que no fim da aula ia ler as faltas, para alguém poder reclamar* ***se { porventura / ?provavelmente }*** *lá estivesse.* (Corpus do Português, Spoken section)
'I remember a class... At the time we were absent and there was an attendant who at the end of the class came to read out the absents, so anyone could protest if perhaps/??probably they were present'

Romanian

(76) *o să fac aşa dacă ?cumva/*probabil va veni Ion*
'let's do it this way if ?perhaps/*probably John comes' (created example)

Having identified a series of non-exclusion of factuality adverbs (adverbs presenting a situation as non-factual without qualifying the commitment of the speaker) for Romance languages, let us examine their distribution within the scope of 'if' complementizers.

While epistemic adverbs cannot modify 'if' complements, non-exclusion of factuality adverbs can; this is illustrated in the examples below with pairs of sentences in each language: 'I didn't understand if perhaps/*probably he arrived'. In these sentences, there is a clear contrast between the first variant with a non-exclusion of factuality adverb ('perhaps': *magari, a lo mejor, porventura*, *cumva* and *des fois*), which is generally quite acceptable, and the second variant with an epistemic adverb ('probably': *probabilmente, probablemente, provavelmente, probabil, probablement*), which is clearly non-grammatical. We tested the difference systematically in our corpora, with pairs of identical sentences, and the results are quite clear-cut, the first variant being always at least ten times more frequent than the second; we then confirmed these results with native speakers.

Italian

(77) *Non ho capito se magari / *probabilmente è arrivato*
'I didn't understand if perhaps/*probably he arrived'

Spanish

(78) *No entendí si a lo mejor /*probablemente llegó*
'I didn't understand if perhaps/*probably he arrived'

Portuguese

(79) *Não percebi se porventura/ *provavelmente chegou*
'I didn't understand if perhaps/*probably he arrived'

Romanian

(80) *N-am inţeles dacă cumva/ *probabil a sosit*
'I didn't understand if perhaps/*probably he arrived'

French

(81) *je n'ai pas compris s'il est arrivé des fois / *probablement*
'I didn't understand if perhaps/*probably he arrived'

If such configurations crop up in speech, they seem invariably linked to deviations from the standard, and can be found for instance in children's productions. In the example below, Jules clearly does not master the epistemic content of *à mon avis* 'in my opinion'.

French

(82) *Je sais pas si maman sera rentrée* ***à mon avis***
'I don't know if mommy will already be home, in my opinion (sic)'
(Jules, 2 years and 10 months)

These results plead for considering Romance 'if' complementizers as NEF markers, i.e. markers of non-exclusion of factuality.

4.4 The value of 'that' complements

We have seen above that the class of predicates taking 'that' complements is quite heterogeneous (it includes many perception, cognition and utterance predicates), and seems not to be specialized for a particular modal meaning.

The complements with 'that' can indeed either represent factual situations endowed with a specific spatio-temporal and referential anchoring:

Italian

(83) *ho visto che c' era un tipo che rideva*
'I saw that there was a guy who laughed'
(> a particular person)

or non-factual situations, i.e., situations that can be interpreted as either referring or non-referring (see above)

Italian

(84) *vorrei che ci fosse un tipo che ride*
'I wish there were a guy who laughs'
(> either a particular person or anybody)

In this sense we might agree with Nordström & Boye's (this volume) claim that 'that' complementizers are epistemically neutral (this volume).

Nevertheless, we think that, at least for Romance languages, this first generalization needs to be refined. It is well known in fact that subjunctive and indicative moods may alternate in Romance 'that' complements. For example, the verbs 'want', 'be strange', 'regret' tend to select the subjunctive mood:

French

(85) *Je voudrais qu' il vienne*
I want.COND.1SG that he come.SBJV.PRS.3SG
'I wish he would come' (come.SBJV.PRS.3SG)

(86) *C' est bizarre qu' il ne vienne pas*
DEM be.IND.PRS.3SG strange that he NEG come.SBJV.PRS.3SG NEG
'It is strange that he does not come' (come.SBJV.PRS.3SG)

(87) *je regrette qu' il ne soit pas venu*
I regret.IND.PRS.1SG that he NEG be.SBJV.PRS.3SG NEG come.PST.PTCP
'I regret he has not come' (be.SBJV.PRS.3SG come.PTCP)

Whereas the verbs 'say', 'see', 'discover' tend to select the indicative mood:

French

(88) *Je vois qu' il est venu*
I see.IND.PRS.1SG that he be.IND.PRS.3SG come.PST.PTCP
'I see that he has come' (be.IND.PRS.3SG come.PTCP)

(89) *Je* *découvre* *qu'* *il* *est* *venu*
I discover.IND.PRS.1SG that he be.IND.PRS.3SG come.PST.PTCP
'I discover that he has come' (be.IND.PRS.3SG come.PTCP)

(90) *J'* *ai* *dit* *qu'* *il* *est* *venu*
I have.IND.PRS.1SG say.PST.PTCP that he be.IND.PRS.3SG come.PST.PTCP
'I said that he has come' (be.IND.PRS.3SG come.PTCP)

This opposition has often been interpreted as an opposition in modality. Hooper (1975), Klein (1975), and Bybee &Terrell (1990) have proposed that the selection of mood is related to a speech-act opposition: according to these authors, asserted situations are in the indicative, while non-asserted situations are in the subjunctive. Palmer (1986) claims that the opposition between indicative and subjunctive has to do with the degree of belief being conveyed: strongly believed situations are in the indicative, weakly believed situations are in the subjunctive. Farkas (1992) and Giannakidou (1994) have argued that the indicative is selected by veridical verbs, i.e., verbs that allow the inference that their complements are taken to be true at least by one epistemic agent (the speaker, or the subject of the main verb); whereas the subjunctive is selected by non-veridical verbs. All these interpretations have advantages as well as shortcomings, as thoroughly illustrated by Marques (2004; 2009). Marques has proposed one of the most recent and convincing interpretations for this opposition. He claims that, at least for Western Romance, the subjunctive is the unmarked choice for the great majority of 'that' complements. We do find, in fact, the subjunctive both within the scope of non-factual verbal predicates,[7]

Italian

(91) *voglio* *che* *sia* *bello*
want.PRS.1SG that be.SBJV.PRS.3SG beautiful.M
'I want him to be beautiful'

and within the scope of factual predicates such as regret:

Italian

(92) *Mi* *dispiace* *che* *sia* *bello*
me.DAT displease.PRS.1SG that be.SBJV.PRS.3SG beautiful.M
'I am sorry he is beautiful'

7 Note that here we are employing the term "factive" in Kiparsky and Kiparsy (1971) and Kartunnen's (1971) sense: a predicate such as regret does not only cast reality over its complement (it is not merely factual) but it also presupposes it as true (it is truly factive).

Besides, we find the subjunctive within the scope of both assertive and non-assertive (presuppositional) verbs[8]:

Italian

(93) *Spero* *che* *sia* *bello*
hope.PRS.1SG that be.SBJV.PRS.3SG beautiful.M
'I hope he is beautiful'

(94) *Sono* *felice* *che* *sia* *bello*
be.PRS.1SG happy that be.SBJV.PRS.3SG beautiful.M
'I'm happy he is beautiful'

On the contrary, the indicative appears in 'that' complements only in one specific semantic situation: when main predicates "express someone's [certainty][9] concerning the truth of their complement proposition" (Marques 2004). Refining Marques's proposal and combining it with Bybee and Terrel's (1990), we claim that such a certainty should be asserted, and not merely presupposed, by the speaker (or by the subject of the sentence). In order to clarify the difference between factual complements, on the one hand, and complements asserted as true, on the other, let us contrast (95) and (96):

Italian

(95) *È* *bello* *che* *sia* *venuto*
be.PRS3SG beautiful that be.SBJV.3SG come.PST.PTCP
'It is great that he has come'

(96) *È* *chiaro* *che* *è* *venuto*
be.PRS.3SG clear that be.PRS.3SG come.PST.PTCP
'It is clear that he has come'

The complements in (95) and (96) are both presented as true, they are both factual. But while in (95) the truth is presupposed, in (96) the truth of the complement is asserted. We can test this by showing that while in (95) the challengeable part of the sentence is the main predicate (97), in (96) the complement is also negotiable (98):

8 By "assertive" verbs we intend, following Simons (2007), verbs that have an asserted complement within their scope.

9 Marques (2004) writes "express someone's belief" rather than certainty, but we prefer to rephrase his formulation in order to avoid ambiguity.

Italian

(97) A: *È bello che sia venuto*
'It is great that he has come'
B: *Non sono d'accordo (non è bello)*
'I don't agree (this is not great)'

(98) A: *È chiaro che è venuto*
'It is clear that he has come'
B: *ma non mi pare proprio* (*che è chiaro* but also *che è venuto*)
'but I really don't think so (I really don't think that is clear and I really don't think that he has come)'

Interestingly, while predicates taking subjunctive 'that' complements cannot have a parenthetical use, predicates taking indicative 'that' complements can have parenthetical uses:

Italian

(99) *Ho visto/scoperto/detto che è rientrato*
'I have seen/ discovered/ said that he has come back'

(100) *è* *rientrato,* *ho* *visto*
be.IND.PRS.3SG come.back.PST.PTCP have.IND.PRS.1SG see.PST.PTCP
'He has come back, I have seen'

(101) *è* *rientrato,* *ho* *scoperto*
be.IND.PRS.3SG come.back.PST.PTCP have.IND.PRS.1SG discovered.PST.PTCP
'He has come back, I discovered'

(102) *è* *rientrato,* *ho* *detto*
be.IND.PRS.3SG come.back.PST.PTCP have.IND.PRS.1SG say.PST.PTCP
'He has come back, I said'

vs.

(103) *Vorrei/ era strano/ mi dispiaceva che fosse rientrato*
'I wish/ it was strange/ I regretted that he had come back'

(104) **fosse* *rientrato,* *vorrei*
be.SBJV.PST.3SG come.back.PST.PTCP want.COND.PRS.1SG
'He has come back, I wish'

(105) **fosse* *rientrato,* *era* *strano*
be.SBJV.PST.3SG come.back.PST.PTCP be.IND.IPFV.3SG strange
'He has come back, it was strange'

(106) **fosse* *rientrato,* *mi* *dispiaceva*
be.SBJV.PST.3SG come.back.PST.PTCP me.DAT dislike.IND.IMPF.3SG
'He has come back, I regretted'

The parenthetical use of a predicate entails in fact that it loses illocutionary autonomy and becomes illocutionarily dependent on its complements (Kahane & Pietrandrea 2012): only asserted (and not presupposed) complements can escape the illocutionary dependency on the main predicate and fulfill the function of nucleus of the illocution (see also Simons 2007, for a similar analysis on asserted factual complements).

We therefore claim that, in spite of the fact that 'that' complements seem not be marked for epistemicity, mood selection in 'that' complements nevertheless leads to a distinction between unmarked 'that' complements and 'that' complements asserted as believed by the speaker: while the former select the subjunctive mood, the latter only select the indicative mood. Table 5 below illustrates this generalization:

Table 5: General distribution of functions between that + subjunctive and that + indicative in Romance

	That + subjunctive	**That +indicative**
Illocutionary status of the complement	*+/– asserted*	*+ asserted*
Factual status of the complement	*+/– factual*	*+ factual*

The analysis of corpus data, though, shows that the picture is still a little bit more complex. Table 6 illustrates the uses of subjunctive and indicative in Western Romance languages. It is clear that the choice of mood is quite oscillating in language use and a diachronic generalization of the indicative seems at play. This situation is quite advanced in French, where the indicative was extended to non-factual asserted complements such as the complements of verbs such as 'think' and 'hope'. We attempt an explanation of this divergence in Section 5.2.

All in all, while it is clear that the 'that' complementizer does not have any modal value by itself, it is possible to identify for some uses, in most Romance languages, a modal value encoded by the construction ['that'+indicative], which conveys asserted belief. From a diachronic point of view, though, there seems to be a trend toward the loss of this modal value, in almost all Western Romance languages. This tendency is at its strongest in French.

Table 6: Verb types and mood of the verb in the subordinate clause (based on corpus data and speaker judgments)

Verb type	French	Italian	Sardinian	Portuguese	Spanish
I say / tell / see / discovered	**(almost always) Indicative**				
I hope / think		*both (more or less equally) available*			
I would like / wish / fear; It's possible / strange	**(almost always) Subjunctive**				

4.5 Distribution of 'if' and 'that' complementizers in Romance: summing up

To sum up what we have discussed in the previous sections, we may say that Romance 'if' and 'that' complements do contrast as to their modal value.

- 'if' complements are always non-factual, namely they represent a specific type of non-factual situations, i.e. non-exclusion of factuality situations. These complements are presented as but one of a set of mutually exclusive options, and speakers do not qualify their epistemic commitment toward the situation, suspending instead their own commitment. In this sense, we claim that non-exclusion of factuality situations depicted by 'if' complements are not modally qualified. For the same reason, we also reject, for Romance, the characterization of 'if' complements as depicting dubitative situations.
- 'that' complements never represent non-exclusion of factuality situations. They may depict either factual or non-factual situations.
  It may be roughly proposed that mood selection plays a role in the modality of 'that' complements, in that indicative 'that' complements tend to express a particular type of modal situation, i.e. a situation asserted as believed by the speaker (or another epistemic agent). Nevertheless, this generalization has to be qualified in light of language use, which shows that the indicative might be drifting toward a generalization of its uses, in the whole Romance area and especially in French.

These general considerations, however, do not mean that 'if' and 'that' complements behave exactly in the same way across Romance languages. There are a number of small differences between languages that we cannot comment on here; we focus in the next section on two specific cases of divergence, which are mostly linked to areal phenomena.

5 The boundaries of the Romance world: The case of Romanian and French

As we noted above, for many verbs introducing complement clauses all Romance languages apply the same pattern, but for other classes of verbs there are slight differences, with either Romanian only or both French and Romanian standing apart from the others – albeit always with distinct phenomena in the two languages. In the two following sections, we focus on differences that can be at least partly explained by taking into account language-external considerations such as the very strong influence of Germanic languages on French, and of Balkan features on Romanian.

5.1 Romanian

The following figures illustrate the distribution of the complementizers *că*, *să* and *dacă* in Romanian. The most striking difference with respect to Western Romance languages is the existence of three complementizers: the first 'that' complementizer, *că*, is selected by perception, cognition and utterance predicates in a similar way to other Romance languages. The 'if' complementizer, *dacă*, is selected by uncertainty predicates and, just like 'if' complementizers in Western Romance, it projects a non-exclusion of factuality meaning on its complements. The second 'that' complementizer, *să*, is selected by predicational predicates, i.e. predi-

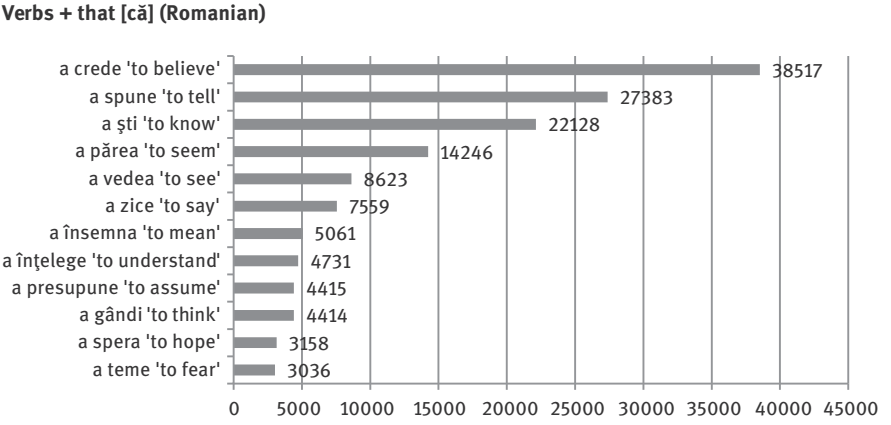


Figure 9: Most frequent verbs introducing complement clauses with 'that$_1$' (*că*) in Romanian

cates which encode "either inception, termination, persistence, success, failure, attempt, intent, obligation or ability – by the subject of the main clause – to perform the action or be in the state that is depicted in the complement clause" (Givón 2001: 150).

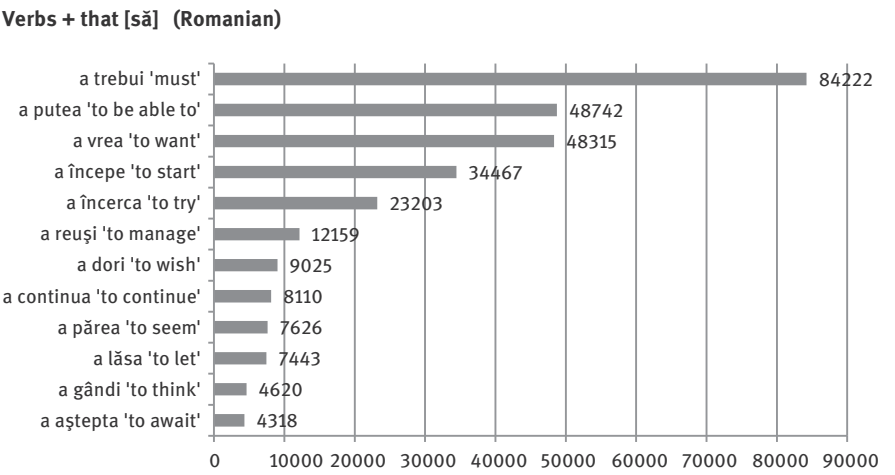


Figure 10: Most frequent verbs introducing complement clauses with 'that$_2$' (*să*) in Romanian

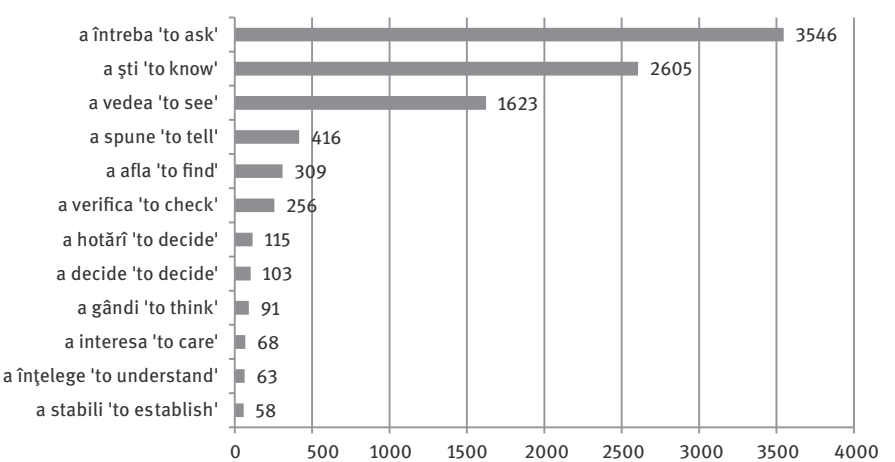


Figure 11: Most frequent verbs introducing complement clauses with 'if' (*dacă*) in Romanian

The differences between Romanian and other Romance languages are not haphazard, but linked to other language-internal features – themselves linked to language-external considerations. As we have seen, the *să* complementizer is used to introduce predicational complements; in these contexts other Romance languages use infinitive clauses (107), although the subjunctive is also possible (108):

French

(107) *Marie a demandé à Jean de lui écrire*
Marie have.PRS.3SG ask.PST.PTCP to Jean COMP DAT.PRO write.INF
'Marie asked John to write to her'

(108) *Marie a demandé à Jean qu' il lui écrive*
Marie have.PRS.3SG ask.PST.PTCP to Jean COMP he DAT.PRO write.SBJV.PRS.3SG
'Marie asked John to write to her'

This specificity of Romanian is linked to a Balkan feature, the weakness of infinitives, i.e. their highly 'nominal' behavior. In Romanian, such utterances are expressed with the subjunctive (i.e. with *să*), to the exclusion of the infinitive (109):

Romanian

(109) *Maria i-a cerut lui Ion să-i scrie*
Maria CL-have.PRS.3SG ask.PST.PTCP DAT.CL Ion COMP-CL write.SBJV.PRS.3SG
'Maria asked Ion to write to her' (Farkas 2003: 2).

Concerning the distribution of indicative and subjunctive in Romanian complements, it should be said that Romanian has a discontinuous subjunctive; the complementizer *să* is thus part of the subjunctive marking (Giannakidou 2009). This is clearly a Balkan feature, present in other languages of the region, including 'outliers' like Serbian (Bulatović 2008). As Table 7 shows, Romanian tends, like French, to generalize the use of the indicative to some optative and epistemic predicates ('hope', 'think', 'be certain'); the *să* complementizer is also used with the subjunctive, for the complements of volitive predicates.

Table 7: Verb types, complementizer and mood of the verb in the subordinate clause, in Romanian (based on corpus data and speaker judgments)

Verb type	Romanian
1. *I know / say / see / tell; I discovered*	*că* + Indicative (also *dacă*)
2. *I hope; I think; It is certain*	
3. *I fear; It's possible / strange*	*că* + Indicative / *să* + Subjunctive
4. *I would like / wish*	*să* + Subjunctive

5.2 French

As we noted for Romanian, the differences between French and other Romance languages are linked to other language-internal features, which are themselves linked to language-external considerations. The main difference observed is the use of the indicative for non-asserted 'that' complements such as those of 'think' and 'hope'. It should be pointed out that this is a recent development: for instance, the verbs *penser* 'to think' and *espérer* 'to hope' could be followed by a 'that' clause in the subjunctive in Old and Middle French, a possibility which gradually disappeared during the 17th and 18th centuries.

This evolution can be explained by the diachronic weakening of the morphological distinction between indicative and subjunctive in French regular verbs. In other Romance languages the choice between subjunctive and indicative remains meaningful in a variety of contexts: the subjunctive mood still has a semantic value. On the contrary, in French, the choice of subjunctive vs. indicative often results from grammatical constraints: for French speakers, more than for instance for speakers of Italian or Spanish, there is an important loss of paradigmatic variability in the use of the subjunctive, as noted by De Mulder and Lamiroy (2012). In this perspective, the subjunctive can be said to have lost (much of) its semantic value in French; in other words, it is more grammaticalized than in 'Central Romance' – in Lehmann's terms, the loss of (here paradigmatic) "variability" is indeed one of the "three main aspects of grammaticalization" (Lehmann 1995: 123).

The fact that the subjunctive, or more precisely the use of the subjunctive in complementation, can be seen as more grammaticalized in French can be brought back to a tendency for French to have, for many constructions, more grammaticalized variants than other Romance languages. This is the case for instance of determiners, with three paradigms in French and only two in almost all other

Romance varieties (Carlier 2007): French has a partitive article which is absent or not completely grammaticalized elsewhere (in Lehmann's terms, the degree of grammaticalization is higher in French on account of paradigmaticity, Lehmann 1995: 123). Similar accounts can be given for demonstratives (Marchello-Nizia 2006: 178), auxiliaries and tense, aspect and modality markers (Lamiroy 1999; 2001; 2003; 2011; De Mulder 2001; De Mulder & Lamiroy 2012); this analysis could even be extended to other areas of language, such as conjunctions, prepositions and discourse markers (Fagard 2010; 2011; Fagard & Mardale 2012).

Of course, this is a global trend and there are counterexamples. However, it is quite pervasive, and could be accounted for by language-external considerations, namely the influence of Germanic languages on French (though not, of course, in a direct way). There is some evidence of a more important impact on the linguistic system at large, and some authors claim that the Germanic superstratum triggered a global phonetic shift in French (Marchello-Nizia 1995: 177). This has been reported quite early on, and Pope (1934: 15) for instance writes that "[t]he Frankish system of accentuation was a strong expiratory one and it was in the intensifying of the weak Latin tonic stress that the Germanic speech-habits, and in particular the Frankish, exercised their strongest influence in pronunciation", with sufficient phonetic consequences (such as diphthongization, reduction of unstressed vowels, closing of some secondary stressed vowels) to explain that these could have resulted in more extended morpho-syntactic changes (see also Von Wartburg [1934] 1965; Zink 1986; for a contradictory analysis, see Noske 2009). This applies to the issue at hand, since the partial loss of contrast between indicative and subjunctive in French is not only semantic, but also formal: in many instances, French verbs have identical forms for both moods while other Romance languages present a formal contrast. An example among others is the form of the verb 'eat': French *mange* (eat.PRS.IND/SBJV.3SG) vs Italian *mangia* (eat.PRS.IND.3SG)/*mangi* (eat.PRS.SBJV.3SG), Catalan *menja* (eat.PRS.IND.3SG)/*mengi* (eat.PRS.SUBJ.3SG), Portuguese and Spanish *come* (eat.PRS.IND.3SG)/*coma* (eat.PRS.SBJV.3SG), Romanian *mănâncă* (eat.PRS.IND.3SG)/*mănânce* (eat.PRS.SBJV.3SG), Sardinian *mandigat* (eat.PRS.IND.3SG)/*mandighet* (eat.PRS.SBJV.3SG), etc.

5.3 A synopsis

To sum up, comparing verb types across Romance languages and their behavior (complementizer + mood) when introducing complement clauses, we identified the following cases:

- Similar pattern across all Romance languages:
  1. Perception, cognition and utterance predicates trigger the use of 'that' + indicative, whatever the language. Uncertainty predicates trigger the use of 'if' + indicative; the subjunctive is also possible.
- Specific pattern for French and Romanian:
  2. Optative verbs and some epistemic verbs ('hope', 'think', 'be certain') trigger the use of 'that' + subjunctive or indicative in most Romance languages, but in French and Romanian the subjunctive is excluded.
- Specific pattern for Romanian only:
  3. Some epistemic verbs as well as verbs of fearing ('be possible', 'be strange', 'fear') trigger the use of 'that' + subjunctive in all Romance languages, except Romanian where *să* + subjunctive alternates with *că* + indicative (see Tables 6 and 7).
  4. Verbs of volition trigger the use of 'that' + subjunctive in all Romance languages, but in Romanian the complementizer is *să* (also + subjunctive) (see Tables 6 and 7).

Table 8 below sums up these findings. Note that some of these are a tendency rather than a rule, and that there is some variation, both regional and diachronic. For instance, as we saw in section Section 5.2., in French verbs of type 2 could introduce complements in the subjunctive until the 18th century.

Table 8: The use of 'that' with indicative and subjunctive and verb types in Romanian and French vs. other Romance languages (Italian, Sardinian, Portuguese, Spanish) (based on corpus data and speaker judgments)

Verb type	'Central' Romance	French	Romanian
1. *I know / say / see / tell; I discovered*	*That + Indicative* (also *If* for the first group)	*That + Indicative* (also *If* for the first group)	*Că + Indicative* (also *dacă* for the first group)
2. *I hope; I think; It is certain*	*That + Subjunctive*		
3. *I fear; It's possible / strange*		*That + Subjunctive*	*Că + Indicative / Să + Subjunctive*
4. *I would like / wish*			*Să + Subjunctive*

6 Origins of Romance complementizers

All standard Romance 'if' forms (except Romanian *dacă*, see below) have the same origin – the 'if' complementizer in Latin, *si* (see i.a. De Dardel 1978). And, as we have seen, they retain its conditional meaning. Note that other constructions can be used to express condition, such as 'when (that)', or even paratactic constructions. Mazzoleni (1997) reports this possibility for various Italian dialects, but it is also found elsewhere (see, for French, Corminboeuf 2008).

Most 'that' forms have the same origin, Latin *quem* 'whom, which, that' (relativizer, in the ccusative), but not Sardinian *ki*, which comes from Latin *qui* 'who, which, that' (also a relativizer, in the nominative). Concerning Romanian, the first 'that' complementizer, *că*, comes from Latin *quod* 'that' (already a complementizer). The second 'that' complementizer, *să*, with more limited uses, comes from the Latin 'if' complementizer (*si*) mentioned above, which remained an 'if' complementizer in all other Romance languages. Actually, it was still used this way in the first literary texts of Old Romanian, but was progressively replaced by *dacă*, which is now the 'if' complementizer in Romanian, and is a compound of the first 'that' complementizer and the adposition *de* 'of, from' (from Latin *dē* 'from (above)').

If Romanian, as we saw above, can be said to diverge from other Romance languages in many respects, there are two main historical reasons. The first is that Romanian has been cut off from the rest of the Romance world for centuries, and thus many innovations which are found in almost all Romance languages did not reach Romanian. This might account for the fact that Romanian kept *quod*, a Late Latin complementizer which was replaced by the relativizer *qui, quem* (Table 3). The second is that Romanian has been heavily influenced by other Balkan languages. This might account for the presence of two paradigms of 'that' complementizers, possibly linked to the fragility of the infinitive and the frequency of use of the subjunctive (of which *să* is considered a part, see Giannakidou 2009). In most Romance languages, 'how' and 'what' complementizers come respectively from Late Latin *quo modo* 'in what way' and *quem* (already seen above); Sardinian *ita* comes from the homonymous Latin adverb meaning 'thus'.

7 Conclusion

In almost all Romance languages, canonical complementizers are restricted to two series, one close to the 'that' complementizer in English, the other to the 'if' complementizer. Romanian presents a different configuration, with three canonical complementizers, two 'that' complementizers and one 'if' complementizer.

Diachronically speaking, most 'that' forms have the same origin, Latin *quem* 'whom, which, that' (relativizer, in the accusative). Almost all 'if' forms have the same origin – the 'if' complementizer in Latin, *si*.

Canonical complementizers can have other uses in Romance besides complementation. 'If' forms are used in hypothetical constructions: this adverbializer use is found in all Romance languages; 'that' forms have a relativizer function in all Romance languages except Romanian and they have a quotative function in French. Both 'if' and 'that' forms can introduce non-assertive main sentences such as jussives, interrogatives and optatives. The combination of complementizers with other subordination markers is a marginal phenomenon; however, it is found in all Romance languages, and could be indicative of a global trend toward a use of the 'that' complementizer as a general subordinator.

From a functional standpoint, Romance complementizers have one main function and an accessory function. The main function is marking the argument predication as illocutionarily dependent on the main predicate (as is proven by the semantic effect of complementizer omission); the accessory function is qualifying the factuality status of the dependent predication. More specifically, 'that' complements, selected by perception, cognition and utterance predicates, may be either factual or non-factual, whereas 'if' complements, selected by uncertainty predicates, are always non-factual.

We proposed that 'if' complements convey a particular type of non-factuality, i.e., non-exclusion of factuality. Speakers present the predication introduced by 'if' as but an eventuality, suspending rather than qualifying their commitment toward its truth. In this sense, we maintained that 'if' complements do not convey epistemic modality.

Concerning 'that' complements, we argued that, at least for Central Romance, mood selection seems to play a role in the modality of 'that' complements, in that indicative 'that' complements tend to express a specific type of modal situation, i.e. a situation asserted as believed by the speaker (or another epistemic agent), whereas subjunctive 'that' complements are unmarked.

Nevertheless, the observation of language use shows that the indicative may be following a trend toward a generalization of its uses in the whole Romance area; this phenomenon is already well under way in French.

Indeed, some areal considerations have led us to refine this general picture. The most important deviation from this generalization is the existence of three complementizers in Romanian: the first 'that' complementizer, *că*, is selected by perception, cognition and utterance predicates; the 'if' complementizer, *dacă*, is selected by uncertainty predicates and, just like 'if' complementizers in Western Romance, it projects a non-exclusion of factuality meaning on its complements; the second 'that' complementizer, *să*, is selected by modality predicates. Modality predicates select infinitive clauses in other Romance languages; the specificity of Romanian is linked to a Balkan feature, the weakness of infinitives.

Concerning French, we have seen that this language seems to be more advanced in the generalization of the usage of indicative in complement clauses. We explained this change with the diachronic weakening of the morphological distinction between indicative and subjective in French regular verbs, which might in turn be partly due to the influence of Germanic languages on French and the tendency of French to have grammatical paradigms which underwent more renewal and differentiation from Latin than in other Romance languages, as has been claimed by several authors.

References

Abeillé, Anne, Danièle Godard & Annie Delaveau. In press. *Grande grammaire du français*. Paris: Actes sud.

Bat-Zeev Shyldkrot, Hava & Suzanne Kemmer. 1988. Le développement sémantique des conjonctions en français : quelques concepts généraux. *Revue Romane* 23(1). 9–20.

Bellert, Irena. 1977. On Semantic and Distributional Properties of Sentential Adverbs. *Linguistic Inquiry* 8(2). 337–351.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow, United Kingdom: Pearson Education.

Blanche-Benveniste, Claire & Dominique Willems. 2007. Un nouveau regard sur les verbes faibles. *Bulletin de la société linguistique de Paris* 102(1). 217–254.

Blanche-Benveniste, Claire. 1989. Constructions verbales en incise et rection faible des verbes. *Recherches sur le français parlé* 9. 53–73.

Blanche-Benveniste, Claire. 1997. *Approches de la langue parlée en français*. Paris: Ophrys.

Boye, Kasper & Peter Harder. 2009. Evidentiality: Linguistic categories and grammaticalization. *Functions of Language* 16(1). 9–43.

Briscoe, Ted. 1994. Prospects for practical parsing of unrestricted text: Robust statistical parsing techniques. In Nelleke Oostdijk & Pieter de Haan (eds.), *Corpus-based research into language*, 67–95. Amsterdam: Rodopi.

Bulatović, Andjelka Gina. 2008. *Modality, Futurity, and Temporal Dependency: the Semantics of the Serbian Perfective Nonpast and Future 2*. Chicago: University of Chicago dissertation.

Bybee, Joan L. & Paul J. Hopper (eds.). 2001. *Frequency and the emergence of linguistic structure*. Amsterdam & Philadelphia: John Benjamins.
Bybee, Joan L. & Tracy D. Terrell. 1990. Análisis semántico del modo en español. In Ignacio Bosque (ed.), *Indicativo y subjuntivo*, 145–163. Madrid: Taurus Universitaria.
Carlier, Anne. 2007. From Preposition to Article. The grammaticalization of the French partitive. *Studies in Language* 31(1). 1–49.
Corminboeuf, Gilles. 2008. 'Tu m'embrasses encore, et c'est mon pied dans les pompons!' Comment construit-on le sens ? *Discours* 3. http://discours.revues.org/4173 (accessed 1 September 2013).
Cornillie, Bert. 2010. On conceptual semantics and discourse functions. The case of Spanish modal adverbs in informal conversation. *Review of Cognitive Linguistics* 8(2). 300–320.
Craig, Colette Grinevald. 1977. *The Structure of Jacaltec.* Austin: University of Texas Press.
Cristofaro, Sonia. 2005. *Subordination.* Cambridge: Cambridge University Press.
Dancygier, Barbara & Eve Sweetser. 2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge: Cambridge University Press.
Dascălu-Jinga, Laurenţia. 2002. *Corpus de română vorbită (CORV) : eșantioane* [Corpus of spoken Romanian (CORV): samples]. București: Editura Oscar Print.
Davies, Mark & Michael J. Ferreira. 2006. *Corpus do Português: 45 million words, 1300s-1900s* [Portuguese Corpus]. Available online at http://www.corpusdoportugues.org.
Davies, Mark. 2002. *Corpus del Español: 100 million words, 1200s-1900s*. Available online at http://www.corpusdelespanol.org.
De Dardel, Robert. 1978. La Forme de la conjonction latine *si* en roman. *Zeitschrift für romanische Philologie* 94. 257–265.
De Mulder, Walter & Béatrice Lamiroy. 2012. Gradualness of grammaticalization in Romance. The position of French, Spanish and Italian. In Kristin Davidse, Tine Breban, Lieselotte Brems & Tanja Mortelmans (eds.), *Grammaticalization and Language Change*, 199–226. Amsterdam & Philadelphia: John Benjamins.
De Mulder, Walter. 2001. La Linguistique Diachronique : grammaticalisation et sémantique du prototype. *Langages* 130. 8–32.
Degand, Liesbeth, Bert Cornillie & Paola Pietrandrea. 2013. Modal Particles and Discourse Markers. Two sides of the same coin? Introduction. In Liesbeth Degand, Bert Cornillie & Paola Pietrandrea (eds.), *Modal Particles and Discourse Markers. Categorization and Description*, 1–18. Amsterdam & Philadelphia: John Benjamins.
Dehé, Nicole & Yordanka Kavalova. 2006. The syntax, pragmatics, and prosody of parenthetical *what*. *English Language and Linguistics* 10. 289–320.
Diewald, Gabriele & Marijana Kresić. 2010. Ein übereinzelsprachliches kontrastives Beschreibungsmodell für Partikelbedeutungen. *Linguistik online* 44. http://www.linguistik-online.de/44_10/diewaldKresic_a.html (accessed 2 July 2013).
Diewald, Gabriele. 2013. "Same same but different" – Modal particles, discourse markers and the art (and purpose) of categorization. In Liesbeth Degand, Bert Cornillie & Paola Pietrandrea (eds.), *Modal Particles and Discourse Markers. Categorization and Description*, 19–25. Amsterdam & Philadelphia: John Benjamins.
Ditchy, Jay K. 1997. *Les acadiens louisianais et leur parler*. Montreal: Comeau & Nadeau.
Elliott, Jennifer R. 2000. Realis and Irrealis: Forms and Concepts of the Grammaticalisation of Reality. *Linguistic Typology* 4. 55–90.
Espinal, Theresa M. 1991. The representation of disjunct constituents. *Language* 67(4). 726–762.

Fagard, Benjamin & Mardale Alexandru. 2012. The pace of grammaticalization and the evolution of prepositional systems: Data from Romance. *Folia Linguistica* 46(2). 303–340.
Fagard, Benjamin. 2010. *É vida, olha…* : Imperatives as discourse markers and grammaticalization paths in Romance: A diachronic corpus study. *Languages in Contrast* 10(2). 245–267.
Fagard, Benjamin. 2011. Conjonctions et grammaticalisation : le cas des langues romanes. *Mémoires de la Société de Linguistique de Paris* 19. 79–102.
Farkas, Donka F. 1992. On the semantics of subjunctive complements. In Paul Hirschbühler & Konrad Koerner (eds.), *Romance Languages and Modern Linguistic Theory*, 71–104. Amsterdam & Philadelphia: John Benjamins.
Farkas, Donka F. 2003. Notes on Mood. Presented at *ESSLLI, Conditional and Unconditional Modality Workshop*, Vienna, 2003.
Frajzyngier, Zygmunt. 1995. A functional theory of complementizers. In Joan Bybee & Suzanne Fleischman (eds.), *Modality in Grammar and Discourse*, 473–502. Amsterdam & Philadelphia: John Benjamins.
Giannakidou, Anastasia. 1994. The semantic licensing of NPIs and the Modern Greek subjunctive. *Language and Cognition. Yearbook of the Research Group for Theoretical and Experimental Linguistics*. 4. Groningen: University of Groningen.
Giannakidou, Anastasia. 2009. The dependency of the subjunctive revisited: temporal semantics and polarity. *Lingua* 120. 1883–1908.
Givón, Talmy. 1982. Evidentiality and Epistemic Space. *Studies in Language* 6(1). 23–45.
Givón, Talmy. 2001 [1984]. *Syntax. A Functional-Typological Introduction*, vol. I. Amsterdam & Philadelphia: John Benjamins.
Haegeman, Liliane. 1988. Parenthetical adverbials: the radical orphanage approach. In Shuki Chiba, Akira Ogawa, Yasuaki Fuiwara, Norio Yamada, Osamu Koma & Takao Yagi (eds.), *Aspects of Modern English Linguistics*, 232–254. Tokyo: Kaitakushi.
Haspelmath, Martin & Ekkehard König. 1998. Concessive conditionals in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial Constructions in the Languages of Europe*, 563–640. Berlin: Mouton de Gruyter.
Herman, József. 1963. *La formation du système roman des conjonctions de subordination*. Berlin: Akademie-Verlag.
Hooper, Joan B. 1975. On assertive predicates. In John P. Kimball (ed.), *Syntax and Semantics* 4, 91–124. New York: Academic Press.
Ifantidou-Trouki, Elly. 1993. Sentential adverbs and relevance. *Lingua* 90. 69–90.
Jakobson, Roman. 1957. Shifters, Verbal Categories, and the Russian Verb. *Russian Language Project*, Department of Slavic Languages and Literatures. Cambridge, MA: Harvard University Press.
Jeanjaquet, Jules. 1894. *Recherches sur l'origine de la conjonction* que *et des formes romanes équivalentes*. Paris: H. Welter, Leipzig: Gustav Fock, Neuchâtel: Attinger.
Kahane, Sylvain & Paola Pietrandrea. 2012. Les parenthétiques comme 'Unités Illocutoires Associées'. Une perspective macrosyntaxique. *Linx* 61(2009). 49–70.
Kahane, Sylvain. 2001. Grammaires de dépendance formelles et théorie Sens-Texte, Tutoriel. *Pr oceedings of TALN'2001*, Tours, France.
Kayne, Richard S. 2005. Some notes on comparative syntax, with special reference to English and French. In Guglielmo Cinque & Richard S. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*, 3–69. New York: Oxford University Press.

Kilgarriff, Adam, Pavel Rychly, Pavel Smrz & D. Tugwell. 2004. The Sketch Engine. In Geoffrey Williams & Sandra Vessier (eds.), *Proceedings of the Eleventh Euralex Congress*, 105–116. Lorient: UBS Lorient.

Kiparsky, Carol & Paul Kiparsky. 1971. Fact. In Danny D. Steinberg & Leon A. Jakobovitz (eds.), *Semantics: an Interdisciplinary Reader in Philosophy, Linguistics and Psychology*, 345–369. Cambridge: Cambridge University Press.

Klein, Flora. 1975. Pragmatic constraints in distribution: the Spanish subjunctive. In *Papers from the IIth regional meeting,* 353–365. Chicago: Chicago Linguistic Society.

Koch, Peter. 2002. Das Französische in typologischer und historisch-vergleichender Sicht. In Ingo Kolboom, Thomas Kotschi & Edward Reichel (eds.), *Handbuch Französisch*, 36–58. Berlin: Schmidt.

Kuno, Susumu. 1973. *The Structure of Japanese Language*. Cambridge, MA & London: MIT Press.

Lamiroy, Béatrice. 1999. Auxiliaires, langues romanes et grammaticalisation. *Langages* 135. 33–45.

Lamiroy, Béatrice. 2001. Le syntagme prépositionnel en français et en espagnol: une question de grammaticalisation? *Langages* 143. 91–106.

Lamiroy, Béatrice. 2003. The External Possession Structure and Grammaticalisation in Romance and Germanic languages. In Yves D'Hulst & Martine Coene (eds.), *From NP to DP. The expression of Possession in Noun Phrases*, Vol. II, 257–281. Amsterdam & Philadelphia: John Benjamins.

Lamiroy, Béatrice. 2011. Degrés de grammaticalisation à travers les langues de même famille. *Bulletin de la Société de Linguistique de Paris* 19. 167–192.

Le Goffic, Pierre. 1993. *Grammaire de la phrase française*. Paris: Hachette.

Lehmann, Christian. 1995. *Thoughts on Grammaticalization*. München & Newcastle: Lincom Europa.

Marchello-Nizia, Christiane, Bernard Combettes, Sophie Prévost & Tobias Scheer (eds.). In preparation. *Grande grammaire historique du français*.

Marchello-Nizia, Christiane. 1995. *L'évolution du français : ordre des mots, démonstratifs, accent tonique*. Paris: Armand Colin.

Marchello-Nizia, Christiane. 2006. *Grammaticalisation et changement linguistique*. Bruxelles: De Boeck Université.

Marques, Rui. 2004. On the system of mood in European and Brazilian Portuguese. *Journal of Portuguese Linguistics* 3(1). 89–109.

Marques, Rui. 2009. On the selection of mood in complement clauses. In Lotte Hogeweg, Helen de Hoop & Andrej Malchukov (eds.), *Cross-Linguistic Semantics of Tense, Aspect, and Modality*, 179–204. Amsterdam & Philadelphia: John Benjamins.

Masini, Francesca & Paola Pietrandrea. 2010. Magari. *Cognitive Linguistics* 21(1). 75–121.

Mauri, Caterina. 2008. The irreality of alternatives: towards a typology of disjunction. *Studies in Language* 32(1). 22–55.

Mazzoleni, Marco. 1997. The syntax of conditional sentences. In Martin Maiden & Mair Parry (eds.), *The Dialects of Italy*, 202–207. Abingdon & New York: Routledge.

Mel'cuk, Igor. 1988. *Dependency syntax: Theory and practice*. Albany: SUNY Press.

Mensching, Guido. 2004. *Einführung in die sardische Sprache*, 3rd edn. Bonn: Romanistischer Verlag.

Mithun, Marianne. 1995. On the Relativity of Irreality. In Joan Bybee & Suzanne Fleischman (eds.), *Modality in Grammar and Discourse*, 367–388. Amsterdam & Philadelphia: John Benjamins.

Noonan, Michael. 2007 [1985]. Complementation. In Timothy Shopen (ed.), *Language Typology and Synctactic Description*, vol 2, *Complex Constructions*, 2nd edn, 52–150. Cambridge: Cambridge University Press.

Nordström, Jackie & Kasper Boye. This volume. Complementizer semantics in the Germanic languages.

Nordström, Jackie. 2010. *Modality and subordinators*. Amsterdam & Philadelphia: John Benjamins.

Noske, Roland. 2009. Autonomous typological prosodic evolution versus the Germanic superstrate in diachronic French phonology. In Enoch Aboh, Elisabeth van der Linden, Josep Quer & Petra Sleeman (eds.), *Romance languages and Linguistic Theory. Selected papers from 'Going Romance' Amsterdam 2007*, 223–242. Amsterdam & Philadelphia: John Benjamins.

Palmer, Frank R. 1987. *Mood and Modality*. Cambridge: Cambridge University Press.

Pietrandrea, Paola. 2011. The conceptual structure of irreality: a focus on non-exclusion-of-factuality as a conceptual and a linguistic category. *Language Sciences* 34. 184–199.

Pope, Mildred.1934. *From Latin to Modern French*. Manchester: Manchester University Press.

Prandi, Michele. 2004. *The Building Blocks of Meaning*. Amsterdam & Philadelphia: John Benjamins.

Ross, John Robert. 1973. Slifting. In Maurice Gross, Morris Halle & Marcel-Paul Schützenberger (eds.), *The Formal Analysis of Natural Language*, 133–169. Berlin: Mouton de Gruyter.

Schelfhout, Carla, Peter-Arno Coppen & Nelleke Oostdijk. 2004. Finite Comment Clauses in Dutch: A Corpus-based Approach. *Journal of Germanic Linguistics* 16. 331–349.

Schoonjans, Steven. 2013. Modal particles: Problems in defining a category. In Liesbeth Degand, Bert Cornillie & Paola Pietrandrea (eds.), *Modal Particles and Discourse Markers. Categorization and Description*, 133–162. Amsterdam & Philadelphia: John Benjamins.

Simons, Mandy. 2007. Observations on embedding verbs, evidentiality, and presupposition. *Lingua* 117(6). 1034–1056.

Squartini, Mario. 2013. From TAM to discourse: The role of information status in North-Western Italian gia 'already'. In Liesbeth Degand, Bert Cornillie & Paola Pietrandrea (eds.), *Modal Particles and Discourse Markers. Categorization and Description*, 163–190. Amsterdam & Philadelphia: John Benjamins.

Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.

Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca: Cornell University Press.

Von Wartburg, Walther. 1965 [1934]. *Evolution et structure de la langue française*, 7th edn. Bern: Francke.

Zink, Gaston. 1986. *Phonétique historique du français*. Paris: Presses Universitaires de France.

Corpora

Romanian: CORV (Dascălu-Jinga 2002), Ruxândoiu (spoken 0.1 million words), Romanian web corpus (written, 44.7 million words).
Spanish: CREA (Real Academia Española, http://www.rae.es), CORDE (Real Academia Española, http://www.rae.es), Corpus del Español (Davies 2002), esTenTen (spoken 7.9 million words, written 2.1 billion words).
Portuguese: Corpus do Português (Davies & Ferreira 2006), ptTenTen11 (spoken 2.2 million words, written 2.8 billion words).
Italian: Badip (De Mauro et al. 1993, accessed via http://badip.uni-graz.at/), itTenTen (spoken 0.5 million words, written 2.6 billion words).
French: Valibel Valibel (Groupe de Recherches Valibel – Discours & Variation, Université catholique de Louvain), PFC (http://www.projet-pfc.net/), Clapi (http://clapi.univ-lyon2.fr), Frantext (Atilf, CNRS, www.frantext.fr), frTenTen (spoken 1.6 million words, written 10.7 billion words).
Note that the Romanian web corpus and the "TenTen" corpora were accessed via the Sketch Engine query system (http://www.sketchengine.co.uk; Kilgariff et al. 2004).

Grammars and dictionaries (selection)

Anglade, Joseph. 1921. *Grammaire de l'ancien provençal*. Paris: Klincksieck.
Avram, Mioara. 1960. *Evoluţia subordonării circumstanţiale cu elemente conjuncţionale în limba română* [Evolution of circumstantial clauses with conjunctive elements in Romanian]. Bucharest: Editura Academiei Republicii Populare Române.
Badia Margarit, Antoni. 1962. *Gramática catalana*. Madrid: Gredos.
Blumenthal, Peter & Achim Stein. 2002. *Tobler-Lommatzsch*, *Elektronische Ausgabe*. Stuttgart: Steiner.
Buridant, Claude. 2000. *Grammaire nouvelle de l'ancien français*. Paris: Sedes.
De Lima Coutinho, Ismael. 1977. *Gramática histórica* [Historical grammar]. Rio de Janeiro: Ao Livro Técnico.
De Mauro, Tullio, Federico Mancini, Massimo Vedovelli & Miriam Voghera. 1993. *Lessico di frequenza dell'italiano parlato*. Milano: Etaslibri.
De Rochegude, Henri Pascal. 1819. *Essai d'un glossaire occitanien pour servir à l'intelligence des poésies des troubadours*. Toulouse: Benichet Cadet.
Densusianu, Ovid. 1938. *Histoire de la langue roumaine*. Paris: Leroux.
Diccionario Medieval Español. 1986. Salamanca: Universidad Pontifica de Salamanca.
Dicţionarul explicativ al limbii române [Explanatory dictionary of Romanian](DEX). 1975. Bucharest: Academia Republicii Socialiste România.
Ditzionàriu de sa limba e de sa cultura sarda [Dictionary of the Sardinian language and culture]. 2000. Cagliari: Condaghes.
Echenique Elizondo, María Teresa & María José Martínez Alcalde. 2003 [2000]. *Diacronía y gramática histórica de la lengua española*, 2nd edn. Valencia: Tirant Lo Blanch.
Ernout, Alfred & François Thomas. 1964. *Syntaxe latine*. Paris: Klincksieck.
Etienne, Eugène. 1980. *Essai de grammaire de l'ancien français (IXe-XIVe siècles)*. Genève: Slatkine Reprints. [Original edition: 1895, Nancy, Paris: Berger-Levrault].

Fabra, Pompeu. 2008 [1933], *Gramàtica catalana* [Catalan Grammar]. Barcelona: Institut d'Estudis Catalans.
Gougenheim, Georges. 1974. *Grammaire de la langue française du seizième siècle*. Paris: Picard.
Gramatica Limbii Române [Romanian Grammar]. 2005. Bucharest: Editura Academiei Române.
Jensen, Frede. 1994. *Syntaxe de l'ancien occitan*. Tübingen: Max Niemeyer.
Koschwitz, Eduard. 1973 [1894]. *Grammaire historique de la langue des Félibres*. Marseille: Laffitte.
Lathrop, Thomas A. 2002 [1984]. *Curso de gramática histórica española*. Madrid: Ariel.
Martorell de Laconi, S. 1990. *Guías de estudio de gramática histórica española*. Salta: Universidad Católica de Salta.
Menéndez Pidal, Ramón. 1929 [1904]. *Manual de gramática histórica española*, 5th edn. Madrid: Victoriano Suarez.
Migliorini, Bruno. 1966. *Storia della lingua italiana*. Firenze: Sansoni.
Moignet, Gérard. 1984 [1973]. *Grammaire de l'ancien français*, 2nd edn. Paris: Klincksieck.
Moll, Francesc de Borja. 1952. *Gramática histórica catalana*. Madrid: Gredos.
Musso, Giancarlo. 2004. *Gramática astësan-a*. Torino: Gioventura Piemontèisa.
Penny, Ralph J. 1993. *Gramática histórica del español*. Barcelona: Ariel Lingüística.
Rohlfs, Gerhard.1954 [1949]. *Grammatica storica della lingua italiana e dei suoi dialetti*. Torino: Einaudi [*Historische Grammatik der italienischen Sprache und ihrer Mundarten*. Bern: Francke].
Tekavčić, Pavao. 1972. *Grammatica storica dell'italiano. II: Morfosintassi*. Bologna: Il Mulino.
Toso, Fiorenzo. 1997. *Grammatica del genovese*. Genova: Le Mani.
Väänänen, Veikko. 1981 [1963]. *Introduction au latin vulgaire*. Paris: Klincksieck.
Vocabolario degli accademici della Crusca. 1866. Accademia della Crusca.
Wagner, Max L. 1950. *La lingua sarda. Storia, spirito e forma*. Bern: Francke.
Zamora Vicente, Alonso. 1985. *Dialectología española*. Madrid: Gredos.
Zauner, Adolph. 1921. *Altspanisches Elementarbuch*. Heidelberg: Carl Winter.

Jackie Nordström and Kasper Boye

Complementizer semantics in the Germanic languages

1 Introduction

All Germanic languages currently spoken have complement clauses. A major distinction can be made between finite and non-finite ones. Non-finite complements are centred around infinitives, participles or nominalizations. They typically designate states-of-affairs, i.e. non-truth-valued meaning units (see Boye 2010; 2012).[1] In contrast, finite complements invariably designate propositions, i.e. truth-valued meaning units. Examples are English *I see that you have a good time* and German *Ich weiß, dass Sie mich hören können* 'I know that you can hear me'. This article deals with finite complement clauses, and more specifically, the complementizers that introduce them.

All Germanic languages display a complementizer contrast in finite complement clauses which is similar to the one in English between *that* (which we call Type 1) on the one hand, and *if* or *whether* (which we call Type 2A and Type 2B, respectively) on the other:

English

(1) a. *He knows* [*that she is there*]
b. *He knows* [*if/whether she is there*]

As will be argued in this chapter, the contrast can be directly linked to truth value: Type 1 complementizers express epistemic neutrality, whereas Type 2 express epistemic uncertainty (see below). We reject the traditional analysis of these elements as speech act markers, Type 1 as declarative and Type 2 as interrogative. Our main arguments against the traditional analysis are firstly that the Type 2 complementizers are found in complement clauses that are not (indirect) questions, and secondly that true subordinate clauses do not have illocutionary force (as demonstrated by Cristofaro 2003).

1 However, there are exceptions. In at least some Germanic languages, for instance, accusative- and nominative-with-infinitive complements construed with an infinitival marker – as in English *I felt him to be a generous spirit* – arguably designate propositions, i.e. truth-valued meaning units (e.g. Dik & Hengeveld 1991: 241–242 on English; Boye 2012: 208–210 on Danish; see Schüle 2000 for a semantic analysis of English complement types).

Type 1 and Type 2 complementizers make up a distributionally delimited morphosyntatic system in the sense of Boye (2012). Thus, both Type 1 and Type 2 can only occur initially in the complement.

English

(2) a. *He knows* [*that/if she is there*].
b. **He knows* [*she is there that/if*].

Furthermore, in the non-colloquial standard varieties of many Germanic languages, Type 1 and Type 2 are mutually exclusive.

English

(3) a. **He knows* [*if that she is there*].
b. **He knows* [*that if she is there*].

Table 1 gives an overview of the canonical finite-clause complementizers in non-colloquial standard varieties of Germanic languages.

Table 1: Canonical finite-clause complementizers in non-colloquial standard varieties of Germanic languages

LANGUAGE	TYPE 1	TYPE 2A	TYPE 2B
Icelandic	*að*		*hvort*
Faroese	*at*	*um*	*hvørt*
Norwegian	*at*	*om*	*hvorvidt*
Swedish	*att*	*om*	*huruvida*
Danish	*at*	*om*	*hvorvidt*
German	*dass*	*ob*	*inwiefern, inwieweit*
Yiddish	*az*	*oyb*	*tsi*
Dutch	*dat*	*of*	
Afrikaans	*dat*	*of*	
Frisian	*dat*	*of*	
English	*that*	*if*	*whether*

As can be seen from the table, all Germanic languages have a Type 1 complementizer, and all have at least one Type 2 complementizer. However, not all have both Type 2A and Type 2B.

The reason for distinguishing between Type 2A and Type 2B has to do with distribution and diachrony. As for distribution, Type 2B complementizers are restricted to complement clauses, whereas in some Germanic languages Type 2A can also occur in adverbial clauses (conditional and hypothetical comparative

clauses; see Section 7 below). This distributional difference can be accounted for diachronically. Type 2A and Type 2B complementizers have different origins (see Section 3 below). Whereas Type 2B complementizers originate from question words within interrogative complement clauses, Type 2A complementizers derive from dubitative nouns that arguably were not part of the complement clauses, but rather functioned as complement-taking predicates. Dubitative nouns would not have been restricted to interrogative contexts and would have had more freedom in developing different uses.

There are other, non-canonical or idiosyncratic finite complementizers in the Germanic languages, such as *how* in English, which can introduce indirect statements etc., and *ifall* in Swedish, which can introduce indirect questions (it otherwise functions as a conditional subordinator, see below). This chapter focuses on Type 1 and Type 2 complementizers, which are found throughout the language family.

The aim of the chapter is two-fold. Firstly, the chapter gives a descriptive overview of the diachrony and grammar of Type 1 and Type 2 complementizers. Thus, after a short presentation of the material on which the chapter is based (Section 2), the paper deals with the diachrony (Section 3), distribution (Section 4), optionality (Section 5), and combinations (Section 6) of the complementizers, and with the non-complementizing uses of complementizer forms (Section 7).

Secondly, as was said above, the chapter argues that the contrast between Type 1 and Type 2 complementizers must be given the following semantic analysis.

> *Semantic analysis of the contrast between Type 1 and Type 2 complementizers*
> Semantically, the contrast between Type 1 and Type 2 complementizers must be understood as a contrast between epistemic neutrality, indeed semantic transparency (Type 1), and uncertainty about the complement proposition (Type 2).

It is argued that all the diachronic and grammatical facts described in sections 3–7 support this analysis. Subsequently, in Section 8, the analysis is contrasted with the standard analysis found in the literature: the analysis of Type 1 and 2 complementizers as indicating assertions and questions respectively. Lastly, we argue against an analysis of the complementizer contrast as one of determined vs. undetermined truth value. Section 9 is a brief conclusion.

2 Material

The data on which this chapter is based is drawn from several sources. Firstly, our own speaker knowledge of the languages in question. This pertains to Danish, Swedish, Norwegian, English and German. Secondly, informants that speak the languages as their mother tongue. Thirdly, reference grammars: Huddleston & Pullum 1997 for English, Haeseryn 1997 for Dutch, Zifonun et al. 1997 for German, Teleman et al. 1999 for Swedish, Faarlund et al. 1997 for Norwegian, and Þráinsson 2005 for Icelandic, among others. Fourthly, dictionaries: *Oxford English Dictionary, Woordenboek der Nederlandsche Taal, Ordbog over det danske Sprog*, *Svenska Akademiens ordbok*, among others. Fifthly, text corpora: *British National Corpus* (*BNC*), *W – Archiv der geschriebenen Sprache* (for German), *KorpusDK* (for Danish), *Korp* (for Swedish), *Oslokorpuset av taggede norske tekster* (for Norwegian) and *Íslenskt textasafn* (for Icelandic).

We have aimed at using authentic data rather than constructed ones. For pedagogical reasons, however, constructed examples are used when presenting e.g. minimal pairs.

3 Diachrony

One reason for grouping Germanic complementizers into three types (1, 2A and 2B) is the fact that all complementizers that belong to the same type are etymologically homogeneous.

Most of the type 1 complementizers probably developed from a demonstrative pronoun (Old English *þæt*, Old High German *daz*, Old Norse *þat*, Gothic *þata*; *Oxford English Dictionary*: "that, conj.") which was part of the matrix clause and referred to a following apposed main clause (Cf. Hopper & Traugott 2003: 190; *Oxford English Dictionary*: "that, conj.").[2] This probable diachrony of Type 1 complementizers is fully compatible with the semantic analysis of them as epistemically neutral: the diachronic source, i.e. demonstrative pronouns, are epistemically neutral themselves.

All Type 2A complementizers seem to derive from a Proto-Germanic subjunction **eƀa*, **ef* (cf. Old English *gif, gyf*, Old High German *obe, oba, ibu*, Old Norse

2 Yiddish *az* derives from Old High German *als* 'as' (Jacobs 2005: 207). Heltoft (1995) suggests that Danish *at* developed from a preposition rather than a pronoun.

ef[3], Gothic *iba, ibai, jabai*), which can ultimately be traced back to a dative form of a noun meaning 'doubt' (cf. Icelandic *ef* and Old Swedish *iäf* 'doubt'; Traugott 1985: 290; Wessén 1956: 220; *Oxford English Dictionary*: "if, conj. and n."; *Woordenboek der Nederlandsche Taal*: "OF (II)"). Traugott (1985) and Wessén (1956) speculate that **eþa, *ef* first functioned as a question word. An even more plausible scenario is that it functioned as a dubitative predicate with the following clause as its complement (or rather apposition) and that it over time grammaticalized into a complementizer (see Nordström 2010: 202–203). Compare the following example from Old Swedish with the noun *iäf*:

Old Swedish

(4) *är thz vtan iäff, [at han hafdhe ey än seet gudh siälwan].*
be.PRS it without doubt COMP he have.PST NEG yet see.PTCP God himself].
'It is without any doubt that he had not yet seen God himself.'
(*Fem Mose Böcker* 1: 91)

As argued in Hopper & Traugott (2003: 187), it is a widespread phenomenon that complementizers develop from predicates of mental states. In further support of the scenario, it should be noted that in some Germanic languages (Dutch, Danish and Norwegian; see below) Type 2A complementizers can be followed by a Type 1 complementizer. This may be indicative of the origin of Type 2A complementizers as predicates.

The origin of Type 2A complementizers as a noun meaning 'doubt' supports the analysis of them as markers of uncertainty. Doubt may be seen as logically equivalent to uncertainty (Boye 2012: 29): Doubt and uncertainty are respectively a positive and a negative construal of approximately the same point of an epistemic scale. Thus, the change in meaning from doubt to uncertainty is natural.

Type 2B complementizers derive from interrogative pronouns. English *whether*, Faroese *hvørt* and Icelandic *hvort* derive from interrogative pronouns meaning 'which of two': Old English *hwæþer*, *hweþer*; Old Norse *hvaðarr* (cf. Old Saxon *hweðar;* Old High German *hwedar*, *wedar*; Gothic *hwaþar*; *Oxford English Dictionary*: "whether, pron., adj. (and n.), and conj."). Danish and Norwegian *hvorvidt* and Swedish *huruvida* have a similar origin, namely Old Norse interrogative pronoun *hwăr* 'which of two' or *hwār* 'which' (of many) + *viðer* 'far' (*Svenska Akademiens ordbok* : "HURU"). German *inwiefern* and *inwieweit* specif-

3 Swedish, Danish and Norwegian *om* derives from Old Swedish, Old Danish and Old Norwegian *um* respectively, and ultimately from *äm, æm, em,* which most likely were sideforms of *ef* (cf. isl. *nefa, nema* 'except if'*Svenska Akademiens ordbok*: "OM.konj.subst2", *Ordbog over det danske Sprog*: "om5 konj."). The development would then have been **eþ > em > um*, (cf. Old Swedish *sem > sum*; rounding of vowel in the presence of a labial consonant; Wessén 1956: 220).

ically derive from prepositional phrases involving interrogative pronouns: *in wie fern* and *in wie weit* ('to what extent'). Yiddish *tsi* is special in that it does not derive from a Germanic interrogative pronoun, but has its origin in Polish *czy* 'if/whether' (Reershemius 2007: 252).

As in the case of Type 2A complementizers, the origin of Type 2B complementizers directly supports the semantic analysis of them in terms of the notion of uncertainty. The interrogative pronouns at hand mark questions which concern a whole proposition, and questions about propositions imply uncertainty about the propositions involved. Thus, the change in meaning from 'question' to 'uncertainty' is natural.

It is remarkable that among the Germanic languages, only Icelandic lacks a Type 2A complementizer. As far as standard Icelandic is concerned, the missing form, *ef*, is used exclusively as a protasis marker in conditional clauses. In Old Icelandic, however, it also functioned as a complementizer, and in young people's colloquial speech, it seems to be regaining this function, as illustrated by the following example (see also Section 7.1):

Old Icelandic

(5) *Ég er næm á tilfinningar fólks og á auðvelt með að*
I be.PRS sensitive on feelings peoples and have.PRS easy with to
sjá [*ef einhverjum líður illa*].
see COMP someone feel.PRS bad
'I'm sensitive to people's feelings and can easily see if somebody feels bad.'
(Almennt blogg á Vefnum)

4 Distribution with complement-taking elements

The distribution of the canonical complementizers is also rather consistent throughout the Germanic languages, although there are certain differences. In all languages, a distinction can be made between 1) complement-taking elements which allow only Type 1 complementizers, 2) complement-taking elements which allow only Type 2 complementizers, and 3) complement-taking elements which allow both Type 1 and Type 2 complementizers. Below, we will first deal with the two restrictive groups (Section 4.1) and then with the non-restrictive one (Section 4.2).

4.1 Restrictive elements

The first group of complement-taking elements, as mentioned, allows only Type 1 complementizers. It is a shared feature of elements belonging to this group that they are semantically incompatible with complement propositions that are uncertain. Either they are non-epistemic and do not imply uncertainty, or they are epistemic and express an epistemic meaning distinct from uncertainty. Icelandic *gerast* ('happen') is an example of a non-epistemic predicate. *Gerast* predicates occurrence of the situation described by its complement proposition, and does not express or imply uncertainty about the proposition at hand. Accordingly, it only allows Type 1 complementizers.

Icelandic

(6) *Það* *gerðist* [*að/*hvort* *sólin* *braust* *fram*].
it happen.PST COMP sun.DEF break.PST forward
'It happened that the sun broke its way through the clouds.'
(Jóhanna Barðdal, p.c.)

English *believe* is an example of an epistemic predicate. *Believe* expresses a positive epistemic stance. It does imply uncertainty about its complement proposition. However, this implication is a scalar implicature which follows from the fact that it *expresses* an epistemic value which ranks below that of 'certainty' on an epistemic scale. The complement proposition is therefore not construed as uncertain. Accordingly, *believe* allows only Type 1 complements.

English

(7) *But Zurich believes* [*that/*if it can make the business profitable with the right package of tailor-made services*].
(The Economist)

In (7), the statement that *Zürich believes [...] it can make the business profitable* does not imply that Zürich is in any way uncertain that it can make the business profitable.

The second group of complement-taking elements allows only Type 2 complementizers. This group includes interrogative utterance predicates like *ask*, predicates of investigation like *investigate*, and predicates of discussion like *discuss*.

English

(8) *Leaving the court, Penn was asked* [*if/*that he wished he had done things differently*].
(Hollywood rogues. Munn, Michael. 1991)

(9) *The long-range effects of mutations in the variable loop we then found for the synthesis of m 2 2G26 prompted us to investigate* [*if/*that also the formation of modifications in the anticodon loop were sensitive to point mutations situated far away in the tRNA*].
(British National Corpus: FTC 857)

Icelandic

(10) *Ég prófaði* [*hvort/*að tölvan virkar*].
I test.PST COMP COMP computer.DEF work.PRS
'I tested whether the computer works.'
(Jóhanna Barðdal, p.c.)

English

(11) *There is debate* [*whether/*that there is an appropriation in such circumstances*].
(British National Corpus: HXE 465)

The complement-taking elements belonging to this second group have in common that they all imply uncertainty about the complement proposition (see Section 3 on the implicational link between questions and uncertainty), and that they do not *express* another distinct epistemic value (for instance, *ask* does not express any epistemic values, but implies uncertainty).

As far as the two restrictive groups of complement-taking elements are concerned, the distribution of complementizers with these elements supports the proposed semantic analysis. On the one hand, the semantic transparency and thus epistemic neutrality of Type 1 complementizers motivates why they are used with a wide range of complement-taking elements. On the other hand, the uncertainty meaning of Type 2 complementizers directly motivates why they are used with elements that imply the meaning of uncertainty: Type 2 complementizers make this meaning explicit. The uncertainty meaning also motivates why Type 2 complementizers are excluded – i.e. why Type 1 complementizers are the only option – with complement-taking elements that are incompatible with uncertain propositions.

4.2 Minimal pairs

The third group of complement-taking elements introduced in the beginning of Section 4 allows both Type 1 and Type 2 complementizers. This group includes assertive utterance predicates like *tell*, perception predicates like *see*, knowledge predicates like *know*, propositional attitude predicates designating certainty

like *certain*, and more marginally propositional attitude predicates designating uncertainty like *uncertain*.

These predicates are all epistemic in the sense that semantically they concern (or can be interpreted as concerning) the source of propositions or the degree of certainty about propositions. In addition, they have in common that (while they may *express* uncertainty, as in the case of *uncertain*) they do not *imply* uncertainty. This distinguishes them from group 1 predicates, which may imply uncertainty (if the implication is due to a distinct epistemic meaning expressed by these predicates), as well as from group 2 predicates, all of which imply uncertainty.

With complement-taking elements belonging to this third group, Type 1 and Type 2 complementizers display a semantic contrast which can be straightforwardly captured by the proposed semantic analysis: Type 2 complementizers mark the complement proposition as uncertain on part of the speaker, while Type 1 complementizers are epistemically neutral. Consider first utterance predicates like the Faroese one in (12).

Faroese

(12) a. *Hon fortelur [at hann fer at koma í dag].*
she tell.PRS COMP he go.PRS to come in day
'She is telling that he will come today.'
(Bergur Rønne Moberg, p.c.)

b. *Hon fortelur [um hann fer at koma í dag].*
she tell.PRS COMP he go.PRS to come in day
'She is telling whether he will come today.'
(Bergur Rønne Moberg, p.c.)

(12b), which contains the Type 2 complementizer *um*, clearly presents the complement proposition ('he will come today') as uncertain. Only *um* can be held responsible for this meaning. Conversely, (12a), with the Type 1 complementizer *að*, clearly does not convey any precise stance towards the complement proposition ('he will come today'). Rather, the epistemic stance conveyed depends on the reliability of the source of the proposition ('she') and on the situational context, including the reliability of the speaker.

Similar things can be said about constructions with knowledge predicates (13), perception predicates (14), and predicates of doubt (15).

Dutch

(13) a. *Hij weet [dat ze thuis is].*
he know.PRS COMP she at.home be.PRS
'He knows that she is at home'.
(Eva van Lier, p.c.)

b. *Hij weet [of ze thuis is*].
he know.PRS COMP she at.home be.PRS
'He knows whether she is at home.'
(Eva van Lier, p.c.)

Icelandic

(14) a. *Hann sér [að sólin braust fram*].
he see.PRS COMP sun.DEF break.PST forward
'He sees that the sun broke through.'
(Jóhanna Barðdal, p.c.)
b. *Hann sér [hvort sólin braust fram*].
he see.PRS COMP sun.DEF break.PST forward
'He sees whether the sun broke through.'
(Jóhanna Barðdal, p.c.)

Danish

(15) a. *Hun tvivler på [at han er der*].
she doubt.PRS on COMP he be.PRS there
'She doubts that he is there.'
b. *Hun tvivler på [om han er der*].
she doubt.PRS on COMP he be.PRS there
'She doubts if he is there.'

In all examples, the b-sentences, which involve Type 2 complementizers, express uncertainty on part of the speaker about the complement proposition, whereas the a-sentences, which involve Type 1 complementizers, do not. The choice between complementizers in this case thus makes possible a distinction between speaker and non-speaker epistemic stance. According to (15b), for instance, the female subject referent doubts the complement proposition 'he is there', and the speaker is uncertain about the proposition. Conversely, (15a), while still attributing doubt about the complement proposition to the female subject referent, may be used to convey, for instance, that the speaker is in fact fully certain about the proposition. Similarly, according to (13a), both the person referred to by the matrix clause subject and the speaker may be certain about the complement proposition 'she is at home'. According to (13b), on the other hand, only the subject referent may be certain; the Type 2 complementizer *of* indicates uncertainty on part of the speaker. Thus, a so-called factive reading of predicates like *weten* 'know' in (13) is conditioned by the complementizers: it is incompatible with the uncertainty meaning of Type 2 complementizers; on the other hand, Type 1 complementizers, being epistemically neutral, are semantically transparent (cf. the semantic transparency of demonstratives, the diachronic sources of Type 1 complementizers), allowing the stance expressed or implied by the complement-taking element to

apply directly to the complement proposition (see Hansen & Heltoft 2011: 1663 on the transparency of Type 1 complementizers in Danish).

Similar contrasts are found also with non-predicative complement-taking elements. The Danish preposition *uanset* ('regardless') is a good example.

Danish

(16) a. *Hun vil være der uanset* [*at han er der*].
she will.PRS be.INF there regardless COMP he be.PRS there
'She will be there even though he is there.'
b. *Hun vil være der uanset* [*om han er der*].
she will.PRS be.INF there regardless COMP he be.PRS there
'She will be there regardless whether he is there.'

The analysis of the Type 2 complementizers as markers of uncertainty is supported by the fact that they are particularly common after predicates of knowledge, perception and saying when these are constructed with future auxiliaries, modal verbs, negation and the polar interrogative, i.e. in irrealis contexts (see Sitta 1971 on German and Nordström 2010: 184 on Swedish):

Swedish

(17) *Då kan jag se* [*om de flyttat isär eller skaffat barn*].
then can I see COMP they move.PTCP apart or get.PTCP children
'Then I will be able to see whether they have moved apart or got children.'
(*Göteborgsposten* 2001)

(18) *Han vet inte* [*om det senare är sant, och ibland tvivlar han*].
He know.PRS not COMP the latter is true and sometimes doubt.PRS he
'He doesn't know if the latter is true, and sometimes he doubts [it]'
(*Dagens nyheter* 1987)

According to Sitta (1971), the common denominator between these kinds of constructions (in German) is "das Moment der Nicht-Festgelegtheit, Unsicherheit".

In fact, Type 2 complementizers are the rule, rather than the exception in these constructions. A corpus investigation of negated predicates of knowledge in Swedish (in *Göteborgsposten* 2011) yielded 268 hits with *vet inte om* 'don't know whether' but only 5 hits with *vet inte att* 'don't know that' (1,8 %). *Att* is used when the speaker, as opposed to the matrix subject, actually knows that the proposition is true:

Swedish

(19) *En hel del vet inte [att sjukdomarna faktiskt kan vara*
a whole part know.PRS not that diseases.DEF actually can be
symptomfria].
symptom.free
'Many don't know that the diseases actually can be free of symptoms.'

Conversely, dubitative predicates (which designate a negative propositional attitude) are only constructed with Type 1 complementizers after negation and rhetorical *wh*-words (since these reverse the negative attitude into a positive one), as the following examples from German show:

German

(20) *Ich zweifle nicht, dass [/*ob] seine Unschuld wieder*
I doubt.PRS not COMP COMP his innocence again
hergestellt wird.
establish.PTCP will
'I don't doubt that his innocence will be established again'
(*St. Galler Tagblatt*, 19 May 2011)

(21) *Wer zweifelt, dass [/*ob] dies geklaut ist???*
who doubt.PRS COMP COMP this steal.PTCP be.PRS
'who doubts that this has been stolen???'
(Alle Wikipedia Diskussionen (Stand: 29.10.2011), A05.52589: Diskussion: Amtsbezeichnung)

To summarize, the distribution of Germanic complementizers supports a semantic analysis according to which Type 1 complementizers are epistemically neutral – in fact, semantically transparent – while Type 2 complementizers indicate uncertainty about the complement proposition. The contrast between the two types of complementizers is obvious with complement-taking elements that allow both types. The contrast also directly motivates that Type 2 is the only possibility with elements that imply (but do not express) uncertainty, whereas Type 1 is the only possibility with many elements that are semantically incompatible with uncertain propositions. Some epistemic complement-taking elements that do not imply uncertainty allow both types of complementizers. With these elements, the contrast between Type 1 and Type 2 complementizers displays a semantic contrast between epistemic neutrality and speaker uncertainty.

As indicated in the beginning of this section, the distribution of the canonical complementizers is not entirely consistent throughout the Germanic languages. Occasionally, complement-taking elements are found that go with Type 1 complementizers in one subset of the languages, and with Type 2 in another. In fact, one

and the same language may possess two complement-talking elements that are semantically closely related, but nevertheless require distinct complementizers. One group of such elements are predicates of pretence. On the one hand, we do not normally pretend what we are certain about. On the other hand, we need to present what we pretend as certain. Accordingly, it is perhaps not entirely unexpected that some predicates of pretence require Type 1 complementizers (e.g. Danish *foregive*), while others require Type 2 (e.g. Danish *lade som om*; cf. e.g. Norwegian *late som om*, Dutch *doen (als)of*, German *tun als ob*). English *pretend* and Swedish *låtsas* even allow both complementizers (*pretend that*; *pretend as if, låtsas att*, *låtsas som om*).

As a final point it must be emphasized that Type 2 complementizers (as well as Type 1 complementizers and all other linguistic expressions) can be used *strategically*. After knowledge predicates, for instance, they can be used in cases where the speaker is actually certain about the complement proposition, but for some reason chooses not to reveal this to the listener:

English

(22) *I greeted him with the age-old aphorism: 'Those who can, do; those who can't, teach; and those who can't teach become administrators – in your case I know whether that is true'.* (*Managing the national curriculum*. ed. Brighouse, Tim & Moon, Bob. 1991)

Some might take this to suggest that Type 2 complementizers do not indicate uncertainty, but rather 'undetermined truth value' (see e.g. Zifonun 1997: 2254 on German and Harder, this volume, on Danish). This would clearly be wrong, however, since other epistemic markers can also be used strategically to withhold information – e.g. *maybe*:

English

(23) *Maybe I have a filing cabinet stuffed under the mat.* (*The waters of Eden*. Neil, Joanna. 1993)

5 Complementizer optionality

Complementizer optionality is often discussed as "complementizer deletion" or "complementizer omission". We prefer to talk about optionality.

Germanic Type 2 complementizers are not optional. Complement-taking elements which allow only Type 2 complementizers (see Section 4 above) do not allow complements without complementizers (cf. Boye et al. 2012: 56 on Danish).

Danish

(24) a. *Hun spørger [om han er der*].
she ask.PRS COMP he be.PRS there
'She is asking if he is there.'
b. **Hun spørger [Ø han er der*].
she ask.PRS he be.PRS there

Type 1 complementizers, conversely, are – in certain contexts – optional in many Germanic languages, but not all. In Danish, English, Faroese, Norwegian, Swedish and Yiddish they are optional in many contexts. The Icelandic Type 1 complementizer *að* can also be optional, but only when it occurs immediately after a finite verb and before a subject pronoun (Thráinsson 1994: 186).

Icelandic

(25) a. *Ég held [að þú megir fullyrða það*].
I think.PRS COMP you may.PRS claim that
'I think that you can claim that.'
b. *Ég held [Ø þú megir fullyrða það*].
I think.PRS you may.PRS claim that
'I think you can claim that.'

In four West Germanic languages, Type 1 complementizers are not (with possible dialect exceptions) optional: Dutch, Afrikaans, Frisian and German. In all these languages, verb-final complements with Type 1 complementizers contrast with verb-second complements without complementizers (Lass 1997: 198–200), but verb-final complements are not found without complementizers.

Frisian (after Lass 1997: 199, citing Sipma 1913: 92)

(26) a. *Hy sei [det er it net verdwaen scoe*]. (verb-final, +COMP)
he say.PRS COMP he it NEG do.again shall.PRS
'He says that he won't do it again.'
b. **Hy sei [Ø er it net verdwaen scoe*]. (verb-final, -COMP)
he say.PRS he it NEG do.again shall.PRS
c. *Hy sei [hy scoe it net verdwaen*]. (verb-second, -COMP)
he say.PRS he shall it NEG do.again
'He says he won't do it again'.

Germanic complementizer optionality raises two questions.

1. Why can Type 1 complementizers be optional, but not Type 2 complementizers?
2. What governs presence vs. absence of Type 1 complementizers?

The semantic analysis advocated here provides a partial answer to the first question. Absence of Type 1 complementizers has no serious semantic consequences: Since Type 1 complementizers are epistemically neutral, the contrast between presence and absence of them is not bound up with any semantic contrast. Conversely, absence of Type 2 complementizers would have serious semantic consequences: Since Type 2 complementizers indicate uncertainty about the complement proposition, the contrast between presence and absence of them would be bound up with a semantic contrast between indicating and not indicating uncertainty about the complement proposition, respectively. It should be emphasized that this answer is only partial. It points out a semantic motivation for optionality of Type 1 complementizers, but does not, for instance, suggest an account of the fact that not all Germanic languages have optional Type 1 complementizers.

As for the second question, a number of factors have been shown or hypothesized to influence the choice between presence and absence of *that* in English, including: type of matrix subject (first and second person singular subjects favour absence), type of complement-taking element (absence is most frequent with non-factive predicates such as *think* and *guess* whereas factive predicates generally favour presence; Platzack 1998: 74; Huddleston & Pullum 2002: 954; Nordström 2010: 210–211), type of complement subject (pronominal subjects favour absence), style (informal style favour absence; e.g. Elsness 1984; Thompson & Mulac 1991; Kaltenböck 2009). Most of these factors can be seen as specific instances of four more general factors that can be assumed to influence on the choice between presence and absence of Type 1 complementizers in those Germanic languages which have optionality (cf. e.g. Boye & Poulsen 2011 on Danish).

One factor is *grammaticalization*. Matrix clauses like English *I think* are known to undergo grammaticalization into particles, grammatical adverbs and even affixes (see e.g. Boye & Harder 2007, 2009). For instance, Afrikaans *ek glo* ('I think') may have developed into an epistemic particle along the lines sketched in (27).

Afrikaans (Thompson & Mulac 1991: 318; Boye 2012: 211)

(27) a. *Ek glo* [*dat hy ryk is*].
I think.PRS COMP he rich be.PRS
'I think that he is rich.'
>

b. *Hy is, glo ek, ryk.*
he be.PRS think.PRS I rich
'He is, I think, rich.'
>
c. *Hy is, glo'k, ryk.*
he be.PRS I.think rich
'He is, I think, rich.'
>
d. *Hy is glo ryk.*
he be.PRS EPIST rich
'He is possibly/presumably/allegedly/seemingly rich.'

A crucial step in this development is the rise in (27b) of clause-internal or -final uses of the original matrix clause in (27a). What is important here, however, is the claim that the matrix clause undergoes grammaticalization also in its clause-initial use (see Boye & Harder 2007, van Bogaert 2011, and the references therein for arguments in support of the claim). This entails a development in which the original complement clause loses its status as complement and subordinate, with the effect that the complementizer becomes superfluous – and even ill-placed – as a complementation marker. Accordingly, absence of Type 1 complementizers is more frequent, the more symptoms or characteristics of grammaticalization are shown by the matrix clause (Boye & Poulsen 2011). For instance, first and second person singular matrix subjects favour grammaticalization, and in English, as mentioned, as well as in Danish they also favour absence of Type 1 complementizers (see Thompson & Mulac 1991 on English, and Boye & Poulsen 2011 on Danish). In Danish, likewise, complement-taking predicates which have a clause-medial use (like Afrikaans *glo ek* in [27b]), favour absence of the Type 1 complementizer *at* (Boye & Poulsen 2011).

The second factor is *disfluency*. Since Type 1 complementizers are optional, speakers can use them as fillers and insert them for rhythmical purposes, or – as pause markers – in order to gain time and hold the floor in case of production difficulties. This has been demonstrated for English *that* by Kaltenböck (2009) (cf. Jaeger 2005 relativizer *that*). In a study of a corpus of spoken Danish, likewise, Boye et al. (2012) found a significant correlation between presence of disfluency symptoms initially in complement clauses and presence of the Type 1 complementizer *at*. Boye et al. also compared the effects of pronominal vs. non-pronominal complement subjects, and found that the latter – which are for different reasons more difficult to process than the former – favour presence of *at* (cf. Elsness 1984 and Thompson & Mulac 1991 on English).

The third factor is *ambiguity*. Contexts in which absence of a Type 1 complementizer results in ambiguity favour presence of the complementizer. Relevant

ambiguities may be of different kinds. In (28) the ambiguity concerns the position of the adverbial *this morning*.

English

(28) *John said this morning the girl was gone.*
(Elsness 1984: 523)

The question is whether the adverbial belongs in the matrix clause (*John said this morning*) or in the complement (*this morning the girl was gone*)? Obviously, presence of *that* resolves the ambiguity. Accordingly, Elsness showed for English that adverbial ambiguities like that in (28) significantly favour presence of *that*.

The fourth factor, which is related to the third one, is *expectation*. Consider first (29).

Danish

(29)	a.	*Jeg*	*er*	*forundret*	*over*	[***det***	***regner***].
		I	be.PRS	puzzle.PTCP	of	**it/that**	**rain.PRS**
		‘I am puzzled that it’s raining.’					
	b.	*Jeg*	*er*	*forundret*	*over*	[***det***	***svar***].
		I	be.PRS	puzzle.PTCP	of	**it/that**	**answer**
		‘I am puzzled about this answer.’					
	c.	*Jeg*	*er*	*forundret*	*over*	[***det***].	
		I	be.PRS	puzzle.PTCP	of	**it/that**	
		‘I am puzzled about it.’					

(29a) involves a temporary ambiguity which resides in the fact that after hearing or reading *det* (‘it/that’), the hearer or reader will not know whether she or he has just heard or read the beginning of a complement clause, as in (29a), or the beginning of a non-clausal complement, as in (29b) and (29c). Presence of the Danish Type 1 complementizer *at* would resolve the ambiguity. It might thus be hypothesized that ambiguous contexts like (29) favour presence of the complementizer. In a study of a corpus of spoken Danish, however, Boye & Poulsen (2011) found only inconclusive support for this hypothesis – perhaps because, prosodic features help resolve ambiguity and thus render *at* superfluous as a disambiguator. On the other hand, it is a remarkable fact that even in languages which have optional Type 1 complementizers, optionality is ruled out in preposed complements.

Danish

(30) a. [*At* *det* *skulle* *være* *mord*] *er* *forkert*.
COMP it should.PST be.INF murder be.PRS wrong
'That it is supposed to be murder is a misunderstanding'.
b. *[*Ø* *det* *skulle* *være* *mord*] *er* *forkert*.
it should.PST be.INF murder be.PRS wrong

The obligatoriness of Type 1 complementizers in cases like (30) may be motivated by the fact that presence of the complementizer rules out the possibility that the complement is interpreted as an independent clause. In a similar vein, adjustment of expectations may explain why high frequency complement-taking elements favour absence of complementizers, and why presence of complementizers seems to be favoured by a long distances between the complement-taking element and the complement clause: English *that* cannot, for instance, readily be left out when there is a longer constituent (phrase) *between that* and the complement-taking element (Huddleston & Pullum 2002: 953, Teleman et al. 1999: 536, Faarlund et al. 1997: 985).

Faarlund et al. (1997: 987) suggest that in Norwegian perception-verb constructions absence of the Type 1 complementizer *at*, as in (31b), is bound up with an evidential reading of the matrix clause.

Norwegian

(31) a. *Jeg* *hører* [*at* *du* *har* *fått* *dig* *nytt*
I hear.PRS COMP you have.PRS become:PTCP yourself new
stereoanlegg].
stereo.system
'I hear that you got yourself a new stereo system (the noise from your flat is louder than before).'
b. *Jeg* *hører* [*Ø* *du* *har* *fått* *dig* *nytt* *stereoanlegg*].
I hear.PRS you have.PRS become:PTCP yourself new stereo.system
'I hear you got yourself a new stereo system (someone told me).'

There is no need to attribute this semantic difference directly to the contrast between presence and absence of *at*. Rather, it can be straightforwardly accounted for in terms of a difference pertaining to grammaticalization. As noted above, absence of Type 1 complementizers is correlated with presence of grammaticalization symptoms. Accordingly, (31b) may be assumed to be relatively strongly grammaticalized. In turn, this would account for the evidential reading of the matrix clause: the matrix clause is developing into a grammatical evidential particle (see Boye & Harder 2007, 2009, 2012).

To sum up, whereas Type 2 complementizers are always obligatory, Type 1 complementizers are – in certain contexts – optional in many Germanic languages, though not all. This pattern supports the semantic analysis advocated in this chapter in that the obligatoriness of Type 2 complementizers may be straightforwardly accounted for as motivated by the fact the these complementizers contribute semantically to the complements, whereas the potential optionality of Type 1 complementizers can be accounted for as motivated by their epistemic neutrality and semantic transparency. In languages and contexts in which Type 1 complementizers are optional, presence vs. absence of complementizers is governed by several factors, including at least 1) the degree of grammaticalization of the matrix clause, 2) presence vs. absence of disfluency around the beginning of the complement, 3) presence vs. absence of ambiguity around the beginning of the complement, 4) the degree with which a clausal complement can be expected.

6 Combinations of canonical complementizers with other complementizers, and with relativizers and adverbial subordinators

One intriguing fact about the canonical complementizers is that in some languages they can be combined. In colloquial English, the two Type 2 complementizers *whether* and *if* are sometimes combined.

English

(32) *The trouble is that no-one appears to know what will happen after the Olympics and* [***whether if*** *by then the United Nations' blacklist will preclude British members of the Tour from competing in certain countries*].
(*Golf Monthly* 1989)

This is also the case in colloquial Swedish and Danish.

Swedish

(33) *Enligt rätten gick det inte att avgöra* [***huruvida om*** *hon fallit eller blivit slagen*].
according.to court.DEF go.PST it NEG to determine **COMP** **COMP** she fall.PTCP or become.PTCP hit.PTCP
'According to the court, it was impossible to determine whether she had fallen or been hit.'
(*Göteborgsposten* 2004)

(32) and (33) both exhibit complementizer combinations of the order 2B–2A. In Danish, the opposite order, 2A–2B, is found more frequently. The reason for this is probably that the Type 2A complementizer is homonymous with the preposition *om* ('about'), which is an obligatory component in complex predicates like *være i tvivl om* ('be in doubt about'), *spørge om* ('ask about'). Thus, due to reanalysis, the combination particle 2B in (34) occasionally gives rise to the combination 2A–2B in (35).

Danish

(34) *Jeg har været i tvivl* ***om*** [***hvorvidt*** *det var*
I have.PRS be.PTCP in doubt **PREP** **COMP** it be.PST
'I have been in doubt as to whether it was the right thing...'
(KorpusDK)

(35) *Årsagen til forbud(d)et er ikke en indikation af* [***om*** ***hvorvidt***
reason.DEF to ban.DEF be.PRS NEG an indication of **COMP** **COMP**
det bliver forbudt at have DeCSS liggende på sin hjemmeside].
it become.PRS prohibited to have DeCSS lie.PTCP on one's homepage
'The reason for the ban is not an indication whether it will be prohibited to put DeCSS on your homepage.'
(KorpusDK)

Combinations of Type 2 complementizers can probably all be dismissed as production errors.

In some languages, however, Type 2 complementizers can also be combined with Type 1 complementizers. This is the case at least in colloquial variants of Dutch, Danish, Norwegian and Icelandic (see e.g. Zwart 1993: 43 on Dutch, and Vangsnes 2006: 3 on Norwegian).

Dutch

(36) *De bondscoach vroeg* [***of*** ***dat*** *het een optie was*]...
DET coach ask.PST **COMP** **COMP** that an option be.PST
'The coach asked if that was an option...'
(www.voetbalzone.nl, 4 November 2006)

Danish

(37) *Reno ved ikke* [***om at*** *han skal have ondt*
Reno know.PRS NEG **COMP COMP** he shall.PRS have sorry
af onkel Arthur].
of uncle Arthur
'Reno doesn't know if he should feel sorry for uncle Arthur.'
(KorpusDK)

(38) *Derefter afklares det* [***hvorvidt at*** *indklagede ønsker*
then clarify.PASS it **COMP COMP** defendant.DEF wish.PRS
dialog].
dialogue
'After that it will be settled whether the defendant wishes to enter a dialogue.'
(www.laeger.dk, 27 September 2012)

Norwegian

(39) *Han er imidlertid spent på* [***om at*** *regjeringen vil*
he be.PRS however excited on **COMP COMP** government.DEF will
innse rederienes vanskelige konkurransesituasjon].
realise shipping.company.PL.DEF.GEN difficult competition.situation.
'He is however excited to see whether the government will realise the difficult situation for the shipping companies.'
(*Bergens tidende* 1995)

Icelandic

(40) *Með nýju kylfurnar getur hann séð* [***hvort að*** *kylfan sé*
With new clubs can.PRS he see.PTCP **COMP COMP** club.DEF be.SBJV
1 eða 2 g þyngri án þess að slá bolta].
1 or 2 grams heavier without to hit ball.INDF
'With new clubs, he can see if the bat is 1 or 2 grams heavier without hitting a ball.'
(*Golfblaðið* 2007)

These combinations, which all have the order Type 2 – Type 1, are much too frequent to be easily dismissed as production errors. They are found both in contexts where Type 2 complementizers are required, as in (36), (38), and (39), and in contexts where Type 2 complementizers contrast with Type 1 complementizers or absence of any complementizer, as in (37) and (40) (see Section 4). One possible account is that these combinations involve Type 1 complementizers used as pause markers (see Section 5). Alternatively, the combinations may be taken to support the hypothesis that, in some languages, Type 1 complementizers are developing into general markers of propositional complements or even of subordination, whereas Type 2 complementizers are epistemic modifiers of the complement propositions.

The latter brings us to the fact that in all Germanic languages, Type 1 complementizers can be added to adverbial subordinators, as in (43) and (44) below (see also Section 7.1 below). Furthermore, in colloquial variants of some Germanic languages, Type 1 complementizers can even be added to relativizers, as in (41) and (42).

Icelandic

(41) *Nú er hluthafi þarna nokkur, Baugur,* [***sem að*** *hefur*
now be.PRS shareholder there some Baugur **REL** **COMP** have.PRS
lýst yfir áhuga á að kaupa stórt fyrirtæki á Bretlandi, Big Food]
declared.PTCP interest on to buy large company on UK.DEF Big Food
'Now there is some shareholder, Baugur, who has declared interest in buying a large company in the UK, Big Food.'
(*Fréttir Ríkissjónvarps* 21–25 September 2004)

Danish

(42) *nogle af de ting* [***som at*** *jeg har prøvet før*] ...
some of those thing **REL** **COMP** I have.PRS try.PTCP before
'some of those things that I have tried before.'
(LANCHART Corpus)

(43) *Det er rart* [***når at*** *det er betalt*].
it be.PRS nice **when** **COMP** it be.PRS pay.PTCP
'It is nice when it has been paid.'
(LANCHART Corpus)

English

(44) *A category A is said to be ... equivalent to a category B* [***provided that*** *there is an equivalence between them*].

In fact, Nordström (2010: 250) shows that more or less every adverbial subordinator in the Scandinavian languages can potentially be combined with a Type 1 complementizer in informal style. In some cases, Type 1 complementizers are optional.

English: *provided (that)*, *so (that)*, *considering (that)*

Dutch: *nu (dat)* 'now', *ingeval (dat)* 'in case', *naarmate (dat)* 'as', *naar gelang (dat)* 'as', *niettegenstaande (dat)* 'notwithstanding the fact that'

Swedish: *efter (det) (att)* 'after', *i det (att)* 'in that', *till dess (att)* 'until', *under det (att)* 'while'

In other cases, they are obligatory.

English: *in order that, in that, in the event that, except that, save that*

Dutch: *voordat* 'before', *eerdat* 'before', *totdat* 'until', *nadat* 'after' *omdat* 'as, because', *zodat* 'so that', *opdat* 'in order that', *doordat* 'by, because', *in plaats (van) dat* 'instead of', *zonder dat* 'without', *behalve dat* 'except that', *zodanig dat* 'such that', *vandaar dat* 'for that reason', *alsmede dat* 'as well as', *inzover(re) dat* 'to the exent that'

German: *so dass* 'so that', *anstatt dass* 'instead of', *ohne dass* 'without', *auf dass* 'on that', *vorausgesetzt dass* 'provided that'

Swedish: *därför att* 'because', *för att* 'in order that', *trots att* 'despite the fact that', *utan att* 'without', *under förutsättning att* 'provided that'

Icelandic: *an þess að* 'without', *fra því að* 'since' (temporal), *med því að* 'because', *af því að* 'since' (cause), *til (þess) að* 'in order to', *sakir þess að* 'because', *úr því að* 'since', *vegna þess að* 'because', *svo að* 'so that', *þannig að* 'so that', *þó að* 'although', *þrátt fyrir það að* 'in spite of the fact that', *eftir að* 'after'

In all cases, the Type 1 complementizer follows the adverbial subordinator. Under the assumption that the relevant Type 1 complementizers are developing into general subordination markers, this ordering is to be expected. Any specification of type or function of subordination – such as the specification of subordinate clauses as adverbial – is dependent upon subordination in the first place. The ordering of Type 1 complementizers may be seen as reflecting this dependency: first, you have a clause; then you mark it as subordinate by adding a Type 1 complementizer (outside the clause at hand); and then you specify the type or function of subordination by adding an adverbial subordinator or a relativizer (outside the Type 1 complementizer). This line of reasoning is supported by the origin of many adverbial subordinators. Adverbial subordinators often derive from constructions consisting of an adverb, preposition, pronoun, participle or prepositional phrase followed by a complement introduced by a Type 1 complementizer (cf. Huddleston & Pullum 2002: 1011–1014). Such constructions are constantly being invented.

English: *despite the fact that, given the fact that*

Dutch: *gezien het feit dat* 'in light of the fact that', *ondanks het feit dat* 'in spite of the fact that' (Donaldson 1997: 230)

Swedish: *genom att* 'by', *mot att* 'in exchange', *i stället för att* 'instead of'.

In any case, the semantic analysis given in the beginning of the present chapter is fully compatible with the combinations in (36)–(44): Since Type 1 complementizers are epistemically neutral, the combination of them with Type 2 complementizers, relativizers or adverbial subordinators does not result in a semantic conflict. In contrast, the combination of Type 2A complementizers with adverbial subordinators is highly restricted. Like Type 1, Type 2A complementizers follow the subordinators they are combined with, but unlike Type 1, they only marginally accompany other subordinators, and crucially – in accordance with their function as uncertainty markers – they are only found in clauses that designate uncertain propositions (concessive conditional and hypothetical comparative clauses; cf. Nordström 2010: 250):

> English: *even if, as if*; Icelandic: *ein og ef* 'as if'; Swedish: *även om* 'even if' *som om* 'as if'; German: *als ob* 'as if'; Dutch: *alsof* 'as if'

In Danish and Norwegian, furthermore, the Type 2A complementizer *om* can be inserted after the conditional protasis marker *hvis* ('if'), as in example (45).

Norwegian

(45)

Ungene	*vokser*	*raskt*	*ut*	*av*	*utstyret*	*og*	*det*	*bliver*	*fryktelig*
children.DEF	grow	quickly	out	of	clothes.DEF	and	it	becomes	terribly
dyrt	[***hvis***	***om***	*man*	*skal*	*kjøpe*	*nytt*	*hvert*	*år*].	
expensive	**CONJ**	**COMP**	one	shall.PRES	buy	new	every	year	

'The children grow right out of their clothes, and it becomes terribly expensive if one must buy new ones each year.'
(*Aftenposten* 1996)

Since all of these clauses describe hypothetical events, *om* can be analysed as indicating uncertainty here as well (see Boye 2008 on Danish and Nordström 2010 on the Germanic languages in general).

It should also be noted that the Type 1 complementizer *at*, due to its epistemical neutrality, is also allowed after *hvis* in Danish and Norwegian, as illustrated in (46).

Danish

(46)

Og	[***hvis***	***at***	*der*	*er*	*nogen*	*i*	*det*	*danske*	*land*	*der*	*vil*
and	**CONJ**	**COMP**	there	be.PRS	anyone	in	the	Danish	country	that	want.PRES
prøve	*at*	*lave*	*denne*	*ret*],	*så*	*siger*	*dronningen:*	*Gud*	*bevare*	*Danmark.*	
try	to	make	this	dish,	so	say	queen.DEF	God	save.SBJV	Denmark	

'And if there is anyone in the Danish country who wants to make this dish, the Queen says: God save Denmark.'
(Danish student assignment 22 May 1992)

To sum up, canonical finite-clause complementizers can in some languages combine with each other or with other subordinators. Most frequently, Type 1 complementizers are found following other subordinators, including Type 2 complementizers. This is compatible with the semantic analysis advocated here: Being epistemically neutral, Type 1 complementizers can be inserted after other subordinators without semantic consequences. Conversely, combinations of Type 2A complementizers with adverbial subordinators are restricted to clauses that designate propositions over which uncertainty is expressed or implied. Finally, combinations of Type 2A and Type 2B complementizers are rather rare and can probably be dismissed as production errors.

7 Complementizers used as other kinds of subordinators

As already Kortmann (1998) pointed out, the Germanic canonical complementizers are notorious for showing up in other clause types. Below, we go through a number of cases in which complementizer forms are found outside complement clauses, in each case relating non-complementizing uses of complementizer forms to the semantic analysis presented in Section 1.

7.1 Adverbial clauses

In many Germanic languages, complementizer forms are used as adverbial subordinators. In certain Germanic languages, Type 1 complementizer forms can – more or less readily – be used to introduce purposive clauses, as in (47). In Danish and other languages, however, a much more common way of introducing purposive clauses is to combine a Type 1 complementizer form with another subordinating element: *så at* ('so that') or *for at* ('in order that') (see Kortmann 1998: 316, referring to Mitchell 1985: 425, 445–452 on Old English; see also Section 6 above).

Danish

(47) *Jeg gjorde det [(for/ så) **at** hun ikke skulle gøre det].*
I do.PST it CONJ/ CONJ **COMP** she not would do it
'I did it so that she didn't have to do it.'

Type 2 complementizer forms are used as the canonical means to introduce conditional protasis clauses in English, Swedish, Faroese and Yiddish[4]:

English

(48) [***If*** *there is no agreement*] *then he will be free to talk to other clubs.*
(British National Corpus)

Swedish

(49) [***Om*** *bilarna kör för fort*], *beror det inte på* [*att*
CONJ cars.DEF drive too fast depend.PRS it not on COMP
hastighetsgränsen är fel].
speed-limit.DEF be.PRS wrong
'If the cars drive too fast, it's not because the speed-limit is wrong.'
(*Dagens nyheter* 1987)

Faroese

(50) *Eg fari* [***um*** *hon kemur*].
I go.PRS **CONJ** he come.PRS
'I will leave, if she comes.'
(Bergur Rønne Moberg, p.c.)

In Danish and Norwegian, the use of Type 2A complementizer forms as protasis markers is more or less obsolete and found mainly in frozen expressions. At least in Danish, however, the canonical protasis marker *hvis* is currently beginning to be used as an alternative to the canonical Type 2 complementizer. Since *hvis* derives from an interrogative pronoun, this may be the beginning of another example of the development of interrogative pronouns into Type 2(B) complementizers (see Section 3 above). In Icelandic and Dutch colloquial speech, likewise, the canonical protasis markers *ef* and *als*, respectively, can be used as alternatives to the canonical Type 2B complementizers *hvort* and *of*, respectively. It is natural to hypothesize that some day Icelandic *ef* will again be conventionalized as a Type 2A complementizer: firstly, Icelandic at present has no Type 2A complementizer; secondly, *ef* is a cognate of Type 2A complementizers in other Germanic languages (see Section 3). In Swedish, finally, not only is the Type 2 complementizer form *om* used as a protasis marker (see above), another protasis marker, *ifall*, is currently showing signs of developing into a Type 2 complementizer. In (51)

4 Yiddish *oyb* is in fact mostly found in conditional protases. At the same time, there are also other conditional subordinators in Yiddish such as *ven* 'when' (Jacobs 2005: 219).

ifall introduces a protasis clause, but in (52) it introduces a regular complement clause.

Swedish

(51) [***Ifall*** *de ändå försöker göra det*] *är det ett brott*.
CONJ they still try.PRS do it be.PRS it a crime
'If they still try to do it, it is a crime.'
(*Dagens Nyheter* 1987)

(52) *Det var osäkerheten om* [***ifall*** *hon skulle hinna fram till sydpolen innan det sista flyget gick från den bemannade amerikanska forskningsstationen den 14 februari*] *som blev avgörande*.
it be.PST uncertainty.DEF about **COMP** she would make.it PRT to South.Pole.DEF before the last plane.DEF leave.PST from the manned American research.station.DEF the 14th February that become.PST decisive
'It was the uncertainty about whether she should make it to the South Pole before the last plane left from the manned American research station the 14 February that became decisive.'
(*Dagens Nyheter* 1987)

Only Dutch and German employ canonical protasis markers that are not formally related to complementizers – *als* or *indien* and *wenn* or *falls*, respectively. This does not disturb the general picture that there is a close relation between Type 2 complementizers and protasis markers.

The semantic analysis presented in Section 1 is perfectly compatible with these non-complementizing uses of complementizer forms. As for the use of Type 1 complementizer forms as purposive clauses, it is compatible with the epistemically neutrality of these forms: purposes do not entail uncertainty or any other specific epistemic stance.

As for the link between Type 2 complementizers and protasis markers, it is captured by the proposed semantic analysis. Like Type 2 complementizers, (at least some) protasis markers can be analysed as indicating uncertainty about the proposition expressed by the clause in their scope. The analysis is as follows: protasis clauses are circumstantial adverbials; they describe circumstances that must obtain for the situation described by the main clause to obtain. Protasis markers add to this an indication that it is uncertain whether the circumstances obtain.

One argument in support of this analysis is that crosslinguistically, polyfunctional expressions are regularly found which can be used either as protasis markers or as uncertainty markers. In Lango (Nilo-Saharan), for instance, the

protasis marker *kónó* 'if' is used outside of protasis clauses in what seems to be harmonic combination with the epistemic adverb *ònò* 'maybe, it is possible'.

Lango

(53) ***Kóno*** *ònò* *àwòt*
if it.is.possible 3SG.go.PFV
'Maybe he's walking.'
(Noonan 1992: 184)

In other languages, the opposite situation is found: general markers of uncertainty (possibility) are used in protasis clauses in harmonic combination with canonical protasis markers (Bybee et al. 1994: 208–210). Furthermore, in Lega-Shabunda (Niger-Congo), the particle *andé* occurs both in protasis clauses (54) and, as a marker of uncertainty, in main clauses (55) and complement clauses (56):

Lega (Botne 1997: 527)

(54) ***Andé*** *mbúla* *zálɔka,* *ntábwɛ.*
EPIST 9.rain 9.?.fall.FIN . NEG.3SG.come.FIN
'Assuming (if) it rains, he won't come.'

(55) ***Andé*** *nźɔka!*
EPIST 9.snake.
'Assume (it's) a snake', 'Presumably (it's) a snake.'

(56) *Nsízi* ***andé*** *ɛ́ndilɛ* *kw* *isɔkɔ.*
1SG.not.know **EPIST** 3SG.go.REC:PST to 5.market.
'I don't know whether/if she went to the marketplace.'

As a second argument, epistemically ambiguous adverbializers are disambiguated as protasis markers in contexts of speaker uncertainty. A case in point is Japanese *-tara*.

Japanese

(57) [*Konya* *syuzin* *ga* *kaette* *ki* *-**tara***], *tazune* *-masyoo.*
tonight husband SBJ returner come **-SUB** ask -will.
'When/if my husband comes home tonight, I'll ask.'
(Akatsuka 1985: 626)

According to Akatsuka (1985: 626), "If the speaker takes for granted that her husband will come home, it is temporal; but if she is not absolutely sure, it is

conditional" (Akatsuka 1985: 626). Similar things could be said about German *wenn* in (58).

German

(58) [*Aber* ***wenn*** *ich* *100* *werde*], *dann* *muss* *ich* *mir* *etwas*
But **CONJ** 1SG.NOM 100 become then must I 1SG.DAT something
einfallen *lassen*.
come.up.with let.
'But when I turn 100/if I live to be 100, then I must come up with something.'
(Mannheimer Morgen, 28.09.2001)

Thus, the use of Type 2 complementizer forms as protasis markers, and the use of protasis markers as Type 2 complementizers are motivated by the fact that both types of expressions indicate uncertainty about the proposition expressed by the clause in their scope. The main difference between the two types of expressions is in the case of the Type 2 complementizers that the conveyed uncertainty applies to argument propositions, while in the case of the protasis markers it applies to circumstantial adjunct propositions.

7.2 Relative clauses

In no Germanic language are Type 2 complementizer forms used as relativizers. In English, however, a cognate of the Type 1 complementizer *that* is used as a relativizer in restrictive relative clauses.

English

(59) *I can't abide a man* [***that*** *can't hold his drink*].
(*Worlds apart*. Cairney, John. 1991)

In Danish colloquial speech, the Type 1 complementizer form *at* is occasionally used to introduce a restrictive relative clause. Unlike what may be the case in English, however, this does not mean that the complementizer form is used as a relativizer. Firstly, the relativizer *som* can always be added to it. Secondly, both in Danish, where the relativizer *som* is optional, and Icelandic, where it is obligatory, the Type 1 complementizer form follows the relativizer (example (41) is here repeated as (61)).

Danish

(60) *Du ligner faktisk lidt en* [***som at*** *jeg kender*] [***som at*** *der (er) sød*].
you look.like actually a.little one **REL** **COMP** I know.PRS **REL** **COMP** there be.PRS nice
'You actually look a bit like someone I know who is nice'.
(Myspace)

Icelandic

(61) *Nú er hluthafi þarna nokkur, Baugur,* [***sem að*** *hefur lýst yfir áhuga á að kaupa stórt fyrirtæki á Bretlandi, Big Food*]
now be.PRS shareholder there some Baugur **REL** **COMP** have.PRS declared.PTCP interest on to buy large company on UK.DEF Big Food
'Now there is some shareholder, Baugur, who has declared interest in buying a large company in the UK, Big Food.'
(*Fréttir Ríkissjónvarps* 21–25 September 2004)

The uses of *at* in (60) and other relative clauses are similar to the uses discussed in Section 7.1. In colloquial Danish, *at* can be inserted more or less freely after almost all kinds of other subordinators (hence, Danish grammarians talk about 'pleonastic *at*').

It is debated whether the English relativizer *that* and the Type 1 complementizer *that* are directly related, or derive independently from the same demonstrative source. As Fischer (1992: 193) demonstrates, the Old English relativizer *þe* was replaced by *that* in Middle English. However, it is unclear whether this *that* was identical to the general subordinator *that*, which could be added to most conjunctions (such as *now that, if that, when that*) or derived independently from the original demonstrative pronoun *that* (Old English *þæt*), which could also be used in relative clauses (see further *Oxford English Dictionary* "that, conj." and "that, pron.2"). If they were directly related, this would be perfectly compatible with the semantic analysis proposed in Section 1. Relative clauses may have a factual implication, as in (62).

English

(62) *Are you the man* [***that's*** *just arrived*]*?*
(*A crowd is not company.* Kee, Robert. 1991)

In uttering (62), the speaker must be certain about the proposition 'the man has just arrived'; otherwise she or he wouldn't use the definite form on the correlate noun (*the man*; see further Hopper & Thompson 1973: 490). Indicating uncertainty, Type 2 complementizer forms would be incompatible with the relative

clause in (62). Type 1 complementizer forms, however, are epistemically neutral and as such compatible with any epistemic stance towards the proposition expressed by the clause in their scope.

8 Germanic complementizer contrasts as epistemic contrasts

We are now in a position to compare the semantic analysis proposed in Section 1 with alternative analyses. Below, we first adduce two additional arguments in support of the analysis of Germanic finite-clause complementizer contrasts as epistemic contrasts (Section 8.1). Subsequently, we argue against two alternative analyses: the analysis of Germanic complementizers as illocutionary expressions (Section 8.2) and as expressions of determined vs. undetermined truth value (Section 8.3).

8.1 Two additional arguments in support of the epistemic analysis

The first additional argument is that many complement-taking elements allow epistemic modal markers that seem to occur in harmonic combination with the Type 2 complementizer.

Danish

(63) *Han spurgte* [***om*** *hun* ***måske*** *var syg*].
he ask.PST **COMP** she **EPIST** be.PST ill
'He asked if she were ill.'

(64) *Hun drikker uanset* [***om*** *han* ***måske*** *gør*].
she drink.PRS irrespective **COMP** he **EPIST** do.PRS
'She is drinking whether or not he is.'

Combinations like those in (63) and (64) clearly support the semantic analysis advocated here. The epistemic modal adverb *måske* 'maybe' expresses a neutral epistemic stance, which is compatible with the uncertainty meaning indicated by the Type 2 complementizer *om*. As opposed to this, epistemic modal adverbs such as Danish *helt sikkert* 'certainly', *sikkert* 'probably' and *sandsynligvis* 'probably', which express a positive epistemic stance not compatible with the uncertainty meaning, do not readily co-occur with Type 2 complementizers. In a corpus

search in a Danish 56 million word corpus, KorpusDK, there were 207 hits with *om* in combination with the epistemic adverbs *muligvis* 'possibly' and *måske* 'maybe', but no hits with *om* in combination with *helt sikkert* 'certainly', *sikkert* 'probably' and *sandsynligvis* 'probably'. The Type 1 complementizer *at*, on the other hand, could be constructed with all of these adverbials (1402 hits for *muligtvis* and *måske* and 224 hits for *helt sikkert*, *sikkert* and *sandsynligvis*; the uneven distribution has to do with the fact that *måske* is one of the most frequent words in Danish.).

The second additional argument in support of the proposed semantic analysis is that complementizers designate epistemic modality in many other languages of the world (see Frajyngier 1994, Noonan [1985] 2007 and Nordström 2010 for numerous examples). Indeed, van Lier & Boye (2010) found that in a random sample of 50 genetically and geographically diverse languages, as many as 13 of the 28 languages that were reported to have complementizers had complementizers with epistemic meanings.

The proposed analysis of Germanic finite-clause complementizers is not without precursors in the literature. For instance, it is related to Nordström's (2010) analysis of the complementizer contrast in terms of realis and irrealis (see also Noonan 2007: 114), as well as to Boye's (2008) analysis of the Danish complementizer contrast as an epistemic modal contrast. However, it is by far not the common analysis. The standard analysis is that the contrast is one of illocutionary force, namely declarative vs. interrogative. This analysis will be refuted in Section 8.2 below. In addition, some researchers argue that complementizer contrast codes a contrast between determined or potential and undetermined truth value. This analysis will be commented on in Section 8.3.

8.2 Arguments against an illocutionary analysis

The standard semantic analysis of Germanic complementizers is that Type 1 and Type 2 complementizers differ in terms of illocutionary force. In a number of studies, Type 1 complementizers are claimed or presupposed to be markers of complements that designate assertions, and Type 2 complementizers to be markers of complements that designate questions (e.g. Huddleston & Pullum 2002 on English, and Teleman et al. 1999 on Swedish). Hence, the two types are often referred to as "declarative" and "interrogative" complementizers respectively.

This analysis is compatible with some of the facts discussed in sections 3–7. Six arguments in support of the analysis seem particularly suggestive. Below, we first discuss these six arguments, in each case arguing that they are at best not

conclusive. Subsequently, we present a major argument against the illocutionary analysis.

The first argument in support of the illocutionary analysis is that complements in which only Type 1 complementizers are possible can undoubtedly be used to report assertive speech acts. This is the case, for instance, when the complement-taking element is an assertive verb.

English

(65) *He claims* [***that*** *it will be a waste of time and money*].
(British National Corpus)

Similarly, complements in which only Type 2 complementizers are possible can undeniably be used to express the content of questions. This is the case, for instance, when the complement-taking element is an interrogative verb, as in (66). But Type 2 complements can be argued to report questions also with other kinds of complement-taking elements – e.g. knowledge verbs, as in (67).

English

(66) *He asked* [***if*** *she was looking forward to the sea voyage*].
(British National Corpus)

(67) *Do you know* [***if*** *Francis had a woman friend*]*?*
(British National Corpus)

Even desubordinated uses of such complements, as in (68) and (69), can be argued to report questions.

German

(68) ***Ob*** *man vielleicht auch mal gewitzte Klassik einladen könnte,*
COMP one maybe also once smart classical.music invite could
ohne sich lustig zu machen?
without REFL fun to make
'Perhaps one could for once invite smart classical music, without making fun of it.'
(Braunschweiger Zeitung, 17 July 2006)

Dutch

(69) ***Of*** *het me smaakt!*
COMP that me taste.PRS
'Oh boy does that taste good!'
(Haeseryn et al. 1997: 548)

This argument is not conclusive, however. The fact, for instance, that Type 2 complements can sometimes be interpreted as reporting questions does not exclude the possibility that they designate uncertain propositions, rather than questions. As already mentioned, polar questions imply uncertain propositions. Thus, in (66) and (67) the polar-question interpretation can straightforwardly be analysed as hinging on the interrogative verb *ask* and the interrogative word order *do you know* respectively. Under this analysis, the complement does not designate a question, but only *what is questioned*, namely *an uncertain proposition*. In a similar vein, (68) and (69) may be analysed (in accordance with Evans 2007) as cases where the question meaning – in the diachronic source of these desubordinated clauses supplied by an interrogative verb – has been conventionally associated with the whole construction, rather than only with the complementizer. What is more, in the majority of cases the illocutionary analysis does not capture intuitions (cf. Zifonun et al. 1997 on German). As discussed earlier, Type 2 complementizers are used with complement-taking predicates designating a whole range of social activities, mental states and other events such as knowledge, perception, saying and discussion. In many cases, analysing the complements of these predicates as designating questions is far-fetched. It is difficult to see, for instance, how the Type 2-complement of a predicate of dependence can be said to designate a question:

English

(70) *Everything depends on* [*whether he appears in the morning.*]
(*The Daily Mirror* 1992)

The second argument in support of the illocutionary analysis is that the analysis of Type 2 complementizers as indicating questions would allow us to give an account of the link between Type 2 complementizers and protasis markers. Just like Type 2 complements, the argument goes, protasis clauses designate questions. One piece of evidence that could be presented in support of this argument is that clauses with interrogative word order can also be used as protasis in some languages:

Danish

(71) [***Går*** ***du***], *går* *jeg* *også*.
go.PRS **you** go.PRS I also
'If you leave, I will leave as well.'

However, this argument is not conclusive either. As discussed in Section 7, there are crosslinguistic arguments that the link between Type 2 complementizers and

protasis markers is the element of uncertainty. In accordance with this, the fact that clauses with interrogative word order can be used as protasis clauses, as in (71), can be seen as motivated by the uncertainty meaning implied by polar questions, rather than by the polar question meaning itself.

Similar objections apply to the third argument in support of the illocutionary analysis. The argument is that the question meaning, which is claimed to be found in Type 2 complementizers is present already in the diachronic source of Type 2B complementizers (see Section 3). One counterargument to this is that the change from interrogative pronouns to Type 2B complementizers need not preserve meaning, but may equally well exploit the link between polar questions and uncertainty: a semantic change from polar question meaning to uncertainty meaning implied by polar question meaning is perfectly conceivable. A second counterargument is that Type 2A complementizers do not derive from interrogative pronouns, but from expressions of doubt, a meaning closely related to uncertainty (see Section 3, Boye 2012: 27–31 and Nordström 2010: 202–203).

The fourth argument is that the illocutionary analysis provides a straightforward account of the fact that Type 2 complements license negative-polarity items such as English *at all* and Danish *noget som helst* ('anything at all').

English

(72) *It would have been opportune for the committee to examine the whole purpose and function of A levels, to see* [***whether*** *they were needed* ***at all***]
(A common policy for education. Warnock, Mary. 1989)

Danish

(73) *Hun* *spørger* [***om*** *han* *har* *lavet* ***noget som helst***].
she ask.PRS **COMP** he have.PRS done.PTCP **anything at all**
'She is asking if he has done anything at all'.
(Boye 2008: 11–12, 20)

Interrogatives are well-known licensors of negative-polarity items (Laduslaw 1996). Adopting the analysis of Type 2 complements as interrogatives, it is perfectly natural that they should license such items as well. Again, the argument is at best not conclusive. Firstly, the uncertainty meaning has a negative component that might equally well license the negative-polarity items. Secondly, negative-polarity items are licensed by a range of markers and contexts that hardly have anything to do with either negative polarity or interrogativity, such as predicates of emotion:

English

(74) *Estabrook retreated from the thieves, lighter in step and cash, but **glad** to be doing so **at all**.*
(Imajica. Barker, Clive. 1992)

(75) *I'm **surprised** [she took such pills **at all**].*
(*A season for murder.* Granger, Ann. 1991)

Here, the proposition is affirmative and presupposed to be true. If anything, it is the modal element of surprise that licences the polarity items in this case. Indeed, in Icelandic, such clauses are obligatorily constructed with the subjunctive form of the verb *skulu* 'shall' (Þráinsson 2005; Sigurðsson 2010). Note that something similar can be said of English:

English

(76) *I'm **surprised** [that she **should** have taken such pills **at all**].*

Thus, negative-polarity items are licenced not only by negative or interrogative elements, but also by modal ones (for a more detailed investigation on this topic, see Giannakidou (1998), who argues that negative polarity items are licenced by non-veridical expressions).

The fifth argument is that the illocutionary analysis accounts for the fact that the polar tag 'or not', so typical for polar interrogatives, can be added to Type 2 complements. Again, however, this does not unequivocally support an analysis of these complements as interrogatives. The possibility of positive-negative alternatives can be interpreted as being licensed by the uncertainty meaning, since full uncertainty is arguably located at the neutral midpoint of a scale which goes from 'full certainty that P' to 'full certainty that not P' (e.g. Boye 2012: 30–31). Note further that 'or not' cannot be added to Type 2 complements after dubitative predicates, since these designate an inclination towards believing that the proposition is not true:

English

(77) *I doubt [if/whether he wrote it *or not].*

This is particularly true of the Type 2A complements. With the Type 2B complements, the situation is actually not as clear-cut as made out by Huddleston & Pullum (2002: 983), since there are quite a few examples in the English and Swedish samples with dubitative nouns and 'or not':

English

(78) *If that is so, the defendant cannot be convicted if he raises doubts as to whether or not he realised that the person with whom he was dealing was a policeman*
(*Offences against public order.* Smith, A. T. H. 1987)

As the examples show, however, the meaning is then no more one of doubt, but rather one of uncertainty.

The sixth argument in support of the illocutionary analysis is that Type 2 complements can at least in Danish be constructed with an interrogative marker that may seem to occur in harmonic combination with the Type 2 complementizer. (79) exemplifies the occurrence of a Danish Type 2 complement with the interrogative particle *mon*.[5]

Danish

(79) *Han spurgte* [***om*** *hun* ***mon*** *var syg*].
he ask.PST **COMP** she **Q** be.PST ill
'He asked if she were ill.'

With many complement-taking elements, however, *mon* is not allowed. This is the case, for instance, with complements of *uanset* 'irrespective'.

Danish

(80) *Hun drikker uanset* [***om*** *han* ***(*mon)*** *gør*].
she drink.PRS irrespective **COMP** he **Q** do.PRS
'She is drinking whether or not he is.'

In contrast, as discussed in Section 8.1, many complement-taking elements permit epistemic modal markers which seem to occur in harmonic combination with the Type 2 complementizer (compare (79) and (80) with (63) and (64) above). This clearly speaks against the illocutionary analysis, and supports the analysis we propose. To be sure, main clause polar questions can also be constructed with adverbs like *måske* and *muligvis* in harmonic combination. This can straightforwardly be attributed to the fact that polar questions imply uncertainty about the proposition they concern. In fact, this is, as far as we can see, the only way to account for these constructions.

A final and major argument against the illocutionary analysis is more theoretical. It has often been argued that dependent clauses do not have illocution-

5 *Mon* derives from the Danish modal verb *monne* 'may' and ultimately from ON *munu* 'may' (*Ordbog over det danske Sprog*: "mon4, adv."). It has a direct parallell in Swedish *månne*.

ary meaning (e.g. Cristofaro 2003; Searle 1969; Nordström 2010: 91–94). If this is correct, complement clauses cannot possibly be analysed as designating assertions and questions, or any other type of speech act. To support this claim, Cristofaro (2003: 37) shows that *that*-clauses do not pass assertiveness tests. They cannot be constructed with tag questions, sentential negation, sentential questions or, as indirect speech, speech-act adverbials (the last example is ours).

English

(81) a. *He said it's raining, *isn't it?*
b. *Is it the case that he said it's raining?* [= 'Did he say it's raining?'/*'Is it raining?']
c. *He said that *frankly/*honestly it was raining*.

Similar things can be said about indirect questions. In e.g. American English, the speech act tag 'or what' can be added to direct questions in order to express frustration or ask for confirmation. However, it cannot be added to indirect questions, especially not in the 3rd person past tense, i.e. when the clause is removed from the present speech-situation:

English

(82) a. *Are you listening, or what?*
b. *He asked if she was listening, *or what*.

8.3 Arguments against the notion of undetermined truth value

As pointed out in Section 4, Zifonun et al. (1997: 2254) argue that the complementizer contrast in German is one of determined vs. undetermined truth value. In a similar fashion, Harder (this volume) argues that the contrast in Danish codes a contrast between potential fact and undecided truth value (see also Ransom 1986: 87–91 who argues that Type 2 complementizers differ from Type 1 in that they express "indeterminate modality"). Their arguments will be treated in what follows.

Both Harder and Zifonun et al. claim that strategic uses of Type 2 complementizers (see example [22] above) do not indicate uncertainty on part of the speaker but rather a withholding of information. However, as shown in example (23) the same can be said about other epistemic markers, such as *maybe*. This does not mean that the primary meaning of *maybe* is not to indicate speaker uncertainty.

A central argument of Harder's against the epistemic analysis is that the uses of Type 2A complementizer forms as protasis markers have more meaning than

mere uncertainty. According to Harder, they pose conditions for the commitment of the speaker to the speech act of the matrix clause. On the other hand, as shown in Section 6, Type 2A complementizers can also follow adverbial subordinators, including pure conditional ones, such as Danish and Norwegian *hvis* (see example [45] above). In these cases, the complementizers are not responsible for the adverbial meaning, but merely that of uncertainty. Furthermore, the fact that the Type 2A complementizers have grammaticalized into conditional subordinators in some of the Germanic languages does not entail that they do not have uncertainty as their core meaning in these cases. In our view, Harder's proposal fails to distinguish two components in the meaning of protasis markers: uncertainty and circumstantiality. Protasis markers are like Type 2 complementizer forms in that they indicate uncertainty about a proposition. They differ from complementizer forms in that, rather than introducing an argument proposition, they introduce a proposition that describe the circumstances under which another proposition (the apodosis proposition) can be true. Because Harder's proposal does not distinguish these two meaning components, it misses the link between protasis markers and Type 2 complementizers: epistemic uncertainty.

Accordingly, the clearest indicator that the Type 2 complementizer forms do *not* indicate undetermined or undecided truth value in themselves but rather indicate uncertainty (apart from additional adverbial meanings) comes from protasis clauses. When protasis clauses are constructed with the polar tag 'or not', which emphasizes that the proposition in this case is truly undetermined, they lose their conditional meaning:

English

(83) a. [*If you come*], *I'll go.*
b. [*If you come or not*], *I'll go.*

Furthermore, as was pointed out in Section 8.2 above, true dubitative constructions with Type 2 complementizers cannot be constructed with 'or not' (see example 77). An analysis of Type 2 complementizers as indicators of undetermined truth value cannot account for this fact.

9 Conclusion

In Section 1, we proposed the following semantic analysis of Germanic finite-clause complementizers:

Semantic analysis of the contrast between Type 1 and Type 2 complementizers
Semantically, the contrast between Type 1 and Type 2 complementizers must be understood as a contrast between epistemic neutrality, in fact semantic transparency (Type 1), and uncertainty about the complement proposition (Type 2).

In sections 3–7, we discussed a number of facts pertaining to those complementizers, in each case arguing that the relevant facts support this analysis.

Diachrony (Section 3)
The analysis of *Type 1* complementizers as epistemically neutral is compatible with the fact that the most likely diachronic sources of these complementizers are cataphoric pronouns, which are epistemically neutral themselves.
The analysis of *Type 2A* complementizers as markers of uncertainty is supported by the fact that these complementizers most likely derive from a noun meaning 'doubt': doubt is logically equivalent to uncertainty, and the change in meaning from doubt to uncertainty is thus natural.
The analysis of *Type 2B* complementizers as markers of uncertainty is supported by the fact these complementizers derive from interrogative pronouns: interrogative pronouns mark questions, and questions imply uncertainty. Thus, the change in meaning from 'question' to 'uncertainty' is natural.

Distribution (Section 4)
The analysis of Type 1 complementizers as epistemically neutral and of Type 2 complementizers as markers of uncertainty is supported by the following facts:
1) *Type 1* complementizers are found in complements of a wide variety of complement-taking elements. 2) *Type 1* complementizers are the only option with complement-taking elements that are incompatible with uncertain complement propositions. 3) *Type 2* complements are the only option with complement-taking predicates that imply uncertainty about their complement proposition.

Optionality (Section 5)
The analysis of *Type 1* complementizers as epistemically neutral is supported by the fact that in many Germanic languages they are optional in certain contexts: the potential optionality of Type 1 complementizers can be accounted for as motivated by their epistemic neutrality and semantic transparency – they do not contribute semantically to the complement and can therefore be dispensed with.
The analysis of *Type 2* complementizers as markers of uncertainty is supported by the fact that they are always obligatory: the obligatoriness of Type 2 complementizers may be straightforwardly accounted for as motivated by the fact that these complementizers contribute semantically to the complements,

Combinations of complementizers (Section 6)
The analysis of *Type 1* complementizers is supported by the fact that it is not uncommon to find these complementizers in combination with other complementizers: Being epistemically neutral, Type 1 complementizers can be inserted initially in dependent clauses without semantic consequences.

The analysis of *Type 2* complementizers as markers of uncertainty is compatible with the fact that combinations of Type 2 complementizers with other subordinators are restricted to complements which designate uncertain propositions.

Non-complementizing uses (Section 7)
The analysis of *Type 1* complementizers as epistemically neutral is compatible with the fact that Type 1 complementizer forms are used as purposive subordinators in certain Germanic languages, and as relativizers in English: purposive and relative clauses do not specifically designate uncertain propositions. The analysis of *Type 2* complementizers as markers of uncertainty is supported by the fact that Type 2 complementizer forms are used as protasis markers in many Germanic languages: like Type 2 complementizers, protasis markers can be analysed as uncertainty markers, and the link between the two kinds of expressions can be seen as motivated by the fact that both indicate uncertainty.

In Section 8, we first adduced a final argument in support of the proposed semantic analysis, namely that complementizers express epistemic modality in many other languages of the world. Subsequently, we argued against what may be considered the standard analysis of Germanic complementizers, namely that they are speech act markers, declarative and interrogative, respectively. Firstly, we pointed out that Type 2 complementizers can introduce clauses that are not indirect questions. Secondly, we argued, in accordance with Cristofaro (2003), Searle (1969) and Nordström (2010) that true subordinate clauses do not qualify as speech acts.

Acknowledgment

We are grateful to Jóhanna Barðdal, Eva van Lier and Bergur Rønne Moberg for helping us with Icelandic, Dutch and Faroese issues, and to Petar Kehayov for valuable comments on a draft of this paper.

References

Corpora used

British National Corpus. 2010. University of Oxford. http://www.natcorp.ox.ac.uk/
Íslenskt textasafn. Ordabók Háskólans. http://www.lexis.hi.is/corpus/leit.pl
Korp. Göteborgs universitet. http://g3.spraakdata.gu.se/korp/
KorpusDK. Det Danske Sprog- og Litteraturselskab. http://ordnet.dk/korpusdk/

Oslokorpuset av taggede norske tekster. Institutt for lingvistiske og nordiske studier. http://www.tekstlab.uio.no/norsk/bokmaal/index.html

W – Archiv der geschriebenen Sprache. Institut für Deutsche Sprache. https://cosmas2.ids-mannheim.de/cosmas2-web/

Literature

Akatsuka, Noriko. 1985. Conditionals and the epistemic scale. *Language* 61 (3). 625–639.

Botne, Robert. 1997. Evidentiality and epistemic modality in Lega. *Studies in language* 21 (3). 509–532.

Boye, Kasper. 2008. Om *om* [About *om*]. *Ny forskning i grammatik* 15. 7–29.

Boye, Kasper. 2010. Reference and clausal perception-verb complements. *Linguistics* 48 (2). 391–430.

Boye, Kasper. 2012. *Epistemic meaning: A crosslinguistic study and a functional-cognitive analysis*. Berlin: Mouton de Gruyter.

Boye, Kasper & Peter Harder. 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in language* 31 (3). 569–606.

Boye, Kasper & Peter Harder. 2009. Evidentiality: Linguistic categories and grammaticalization. *Evidentiality in language and cognition*. [Special issue]. *Functions of language* 16 (1). 9–43.

Boye, Kasper & Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88 (1). 1–44.

Boye, Kasper & Mads Poulsen. 2011. Complementizer deletion in spoken Danish. Paper presented at Societas Linguistica Europaea – 44th Annual Meeting, Universidad de la Rioja, 8–11 September.

Boye, Kasper, Mads Poulsen, Hanna Bruun Pedersen, Marie Herget Christensen, Line Dalberg & Nicoline Munck Vinther. 2012. At eller ikke at i tale og skrift [*At* or not *at* in speech and writing]. *NyS* 42. 41–61.

Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Carlson, Robert. 1994. *A Grammar of Supyire*. Berlin: Mouton de Gruyter.

Cristofaro, Sonia 2003. *Subordination.* Cambridge: Cambridge University Press.

De Schutter, Georges. 1994. Dutch. In Ekkehard König & Johan van der Auwera (eds.), *The Germanic languages,* 439–477. London: Routledge.

Donaldson, Bruce. 1997. *Dutch: A comprehensive grammar*. London: Routledge.

Elsness, J. 1984. *That* or zero? A look at the choice of object clause connective in a corpus of American English. *English Studies* 65. 519–33.

Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 366–431. Oxford: Oxford University Press.

Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk* [Norwegian reference grammar]. Oslo: Universitetsforlaget.

Fischer, Olga. 1992. Syntax. In Rochard M. Hogg (ed.), *The Cambridge history of the English language*. Cambridge: Cambridge University Press.

Frajzyngier, Zygmunt. 1995. A functional theory of complementizers. In Joan Bybee & Suzanne Fleischman (eds.), *Modality in grammar and discourse,* 473–502. Amsterdam: Benjamins.

Giannakidou, Anastasia. 1998. *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins.
Haberland, Hartmut. 1994. Danish. In Ekkehard König & Johan van der Auwera (eds.), *The Germanic languages*, 313–348. London: Routledge.
Haeseryn, Walter (ed.) 1997. *Algemene Nederlandse spraakkunst* [General Dutch grammar]. Groningen: Nijhoff.
Hansen, Erik & Lars Heltoft. 2011. *Grammatik over det danske sprog* [Grammar of the Danish language], vol. I–III. Copenhagen: Det Danske Sprog- og Litteraturselskab.
Harder, Peter. This volume. Complementizers and the complementizing function: Structure and pragmatics in a panchronic perspective.
Heltoft, Lars. 1995. Grammatikaliseringsprocesser i dansk syntakshistorie [Grammaticalization processes in the syntactical history of Danish]. In Bente Holmberg (ed.), *Sproghistorie i 90'erne*, 125–168. Copenhagen: Selskab for Nordisk Filologi.
Hopper, Paul & Elizabeth Closs Traugott. 2003. *Grammaticalization*, 2nd edn. Cambridge: Cambridge University Press.
Huddleston, Rodney D. & Geoffrey K. Pullum (eds.). 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
Jacobs, Neil G. 2005. *Yiddish: A linguistic introduction*. Cambridge: Cambridge University Press.
Jaeger, T. Florian. 2005. Optional *that* indicates production difficulty: Evidence from disfluencies. *Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop*, Aix-en-Provence, France, 10–12 September 2005.
Kaltenböck, Gunther. 2009. Is *that* a filler? On complementizer use in spoken object clauses. *Vienna English working papers* 18 (1). 28–63.
Kortmann, Bernd. 1998. Adverbial subordinators in the Languages of Europe. In Johan van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 457–562. Berlin: Mouton de Gruyter.
Ladusaw, W. 1996. Negation and polarity items. In Shalom Lappin (ed.), *The Handbook of contemporary semantic theory*, 321–342. Oxford: Basil Blackwell.
Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.
Mitchell, B. 1985. *Old English syntax*. Oxford: Oxford University Press.
Noonan, Michael 2007 [1985]. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description, vol. 2, Complex constructions*, 2nd edn., 52–150. Cambridge: Cambridge University Press.
Noonan, Michael P. 1992. *A grammar of Lango*. Berlin: Mouton de Gruyter.
Nordström, Jackie. 2010. *Modality and subordinators*. Amsterdam: John Benjamins.
Ordbog over det danske sprog [Dictionary of the Danish language]. Det Danske Sprog- og Litteraturselskab. http://ordnet.dk/ods/
Oxford English dictionary. 2012. Oxford: Oxford University Press. http://www.oed.com
Platzack, Christer. 1998. A visibility condition for the C-domain. *Working papers in Scandinavian syntax* 61: 53–99.
Ransom, Evelyn N. 1986. *Complementation: Its meaning and forms*. Amsterdam: John Benjamins.
Rersheemius, Gertrud. 2007. Grammatical borrowing in Yiddish. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 245–260. Berlin: Mouton de Gruyter.

Schüle, Susanne. 2000. *Perception verb complements in Akatek, a Mayan language*. Tübingen: University of Tübingen dissertation.
Sigurðsson, Halldór Ármann. 2010. Mood in Icelandic. In Björn Rothstein & Rolf Thieroff (eds.), *Mood in the Languages of Europe*, 33–55. Amsterdam: John Benjamins.
Sitta, H. 1971. *Semanteme und Relationen: Zur Systematik d. Inhaltssatzgefüge im Dt.* Frankfurt: Athenäum-Verlag.
Svenska Akademiens ordbok [The Swedish Academy dictionary]. Göteborgs universitet. http://g3.spraakdata.gu.se/saob/
Teleman, Ulf, Staffan Hellberg & Erik Andersson. 1999. *Svenska Akademiens grammatik* [The Swedish Academy grammar]. Stockholm: Nordsteds Ordbok.
Thompson, Sandra A. & Anthony Mulac. 1991. A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization* 2, 313–329. Amsterdam: John Benjamins.
Thráinsson, Höskuldur. 1994. Icelandic. In Ekkehard König & Johan van der Auwera (eds.). *The Germanic Languages*, 142–189. London: Routledge.
Þráinsson, Hölskuldur. 2005. *Íslensk tunga III: Setningar: Handbók um setningafræði* [The Icelandic language III: Clauses: Handbook of syntax]. Reykjavík: Almenna bókafélagið.
Traugott, Elizabeth Closs. 1985. Conditional markers. In John Haiman (ed.), *Iconicity in syntax,* 289–307. Amsterdam: John Benjamins.
van Bogaert, Julie. 2011. I think and other complement-taking mental predicates: A case of and for constructional grammaticalization. *Linguistics* 49 (2). 295–332.
van Lier, E. & Kasper Boye 2010. Epistemic complementizers: A crosslinguistic survey. Paper presented at Institut für Linguistik, Universität Leipzig, 14 January.
Vangsnes, Øystein. A. 2006. Syntactic doubling phenomena in Scandinavian. *Workshop on syntactic doubling in European dialects*. Amsterdam: Meertens Institute.
Wessén, Elias 1956. *Svensk språkhistoria* III: *Grundlinjer till en historisk syntax* [The history of the Swedish language III: Outline of a historical syntax]. Edsbruk: Akademitryck.
Woordenboek der Nederlandsche taal [Dictionary of Spoken Dutch]. 2007–2010. Instituut voor Nederlandse Lexicologie. http://gtb.inl.nl
Zifonun, Gisela, Ludger Hoffmann, Bruno Strecker & Joachim Ballweg. 1997. *Grammatik der deutschen Sprache*. Berlin: Walter de Gruyter.
Zwart, Jan Wouter. 1993. *Dutch syntax: A minimalist approach*. Groningen: University of Groningen dissertation.

Björn Hansen, Alexander Letuchiy and Izabela Błaszczyk

Complementizers in Slavonic (Russian, Polish, and Bulgarian)

1 Introduction

This contribution consists of three parts. The first is dedicated to a description of the complementizer system in Russian, whereas the second and the third describe the same aspects of Polish and Bulgarian – though not in the same detail as for Russian. In the Russian part, the set of Russian complementizers will be considered, each of them will be described in short, and, finally, some problematic cases of multifunctional units will be also mentioned.

We define a complement clause as a constituent with the internal structure of a clause which functions as a semantic core argument of the matrix verb of the superordinate clause. In a complement clause the core arguments are encoded in the same way as in a main clause (cf. Dixon 2006). Note that, as we say below, the very distinction between arguments and adjuncts, as well as between sentential arguments (complement clauses) and sentential adjuncts is far from being clear.

Though syntactic and morphological strategies of forming clausal complements vary significantly across languages, we follow Dixon's (2006) view and exclude from our analysis all complements lacking tense and mood markers like for example infinitival phrases, though some remarks on infinitives will be included. We define complementizers as a structurally homogeneous subclass of conjunctions linking matrix and complement clauses. The infinitive marker is excluded, because it differs from conjunctions both from morphological and syntactic point of view.

We exclude from the discussion subordinate clauses introduced by a correlative pronoun in the matrix clause, because here the subordinate clause does not function as a core argument of the verb of the superordinate clause; cf. (1):[1]

1 The Russian examples are taken from the National Corpus of Russian (*Nacional'nyj korpus russkogo jazyka*; www.ruscorpora.ru) or from the Internet. According to the concept of the present volume, in some cases we have constructed examples in order to illustrate specific functional contrasts.

(1) *Choč-u* *Vas* *poblagodari-t'* *za* ***t-o***, [***čto*** *Vy*
want-PRS.1SG you thank-INF for **that-ACC.SG.N** **that** you.NOM
otkliknu-l-i-s' *na* *mo-e* *priglašeni-e*].
react-PST-PL-REFL to my-ACC.SG invitation-ACC.SG
lit. 'I would like to thank you for that that you replied to my invitation.'

In this example, the subordinate clause depends on the correlative *to* and not on the verb *poblagodarit'*[2].

Of course, the (un)availability of the correlative is not independent from the verbal semantics. For instance, Kobozeva (2012) and Zaliznjak & Mikaeljan (1986) note that the correlative often occurs in factive contexts. On the other hand, according to Letuchiy (2012) and Zimmerling (2014) correlatives are characteristic for sentential complements with particular semantic roles: for instance, agentive roles (causer / agent) are rarely coded by sentential complements without a correlative.

Consider the following remarks about the structures with correlatives. First of all, the distribution of the variant with and without a correlative is far from being trivial. For instance, with the verb *blagodarit'* 'thank' which hosts a complement with *za* 'for', mainly the variant with correlative is used: if the preposition *za* and the correlative *to* are removed, the phrase sounds colloquial or even stylistically bad, if not ungrammatical. However, this is not true for the intransitive verb *nadejat'sja* 'hope' which also hosts a prepositional complement, but with the preposition *na*. The two variants, with the correlative (*nadejat'sja na to, čto* 'hope that') and without it (*nadejat'sja, čto* 'hope that') are equally in use and none of them seems to be stylistically poor:

(2) *Vs-e* *nadeja-l-i-s'* (***na*** ***t-o***,) [***čto*** *on*
everyone-NOM.PL hope-PST-PL-REFL **on** **that-ACC.SG.N** **that** he.NOM
ostan-et-sja *ži-t'*].
stay-PRS.3SG-REFL live-INF
'Everyone hoped that he will stay alive.'

Second, not only the complementizer *čto*, but also the other complementizers discussed here have variants with correlatives, though only the structure with *to čto* has been discussed in detail in the modern literature (see e.g. Zaliznjak & Mikaelyan 1988, Knjazev 2009, Kobozeva 2011); cf. (3) with *to, kogda*.

2 This type of structure is called "pronominal correlative" (*mestoimenno-sootnositel'nye*) in traditional Russian linguistics, such as Belošapkova (1977) and Valgina (2000).

(3) *Nikak-ie* *dann-ye* *ne* *mog-ut* *da-t'*
neither-NOM.PL data-NOM.PL not can-PRS.3PL give-INF
ukazanie *na* ***t-o*** [***kogda*** *ėto* *proizoš-l-o*].
indication.SG.ACC on **that-ACC.SG.N** **when** this.NOM.SG.N happen-PST-SG.N
'No data can point to the time when it happened.'

Each of these variants has specific conditions of use, but all of the pairs with vs. without a correlative have one general feature: the variant with the correlative is specialized for factive contexts (≈ the event in the embedded clause either has taken place or must take place with no doubt). This is not true for the variant without the correlative. For instance, both the variant with a correlative and without it are possible with the verb *znat'* 'know' and other factive verbs of knowledge. In contrast, non-factive verbs like *dumat'* 'think' and *ščitat'* 'suppose' can only host structures without correlatives.

(4) *Ja* *znaj*-u (***t-o***) [***čto*** *ničego* *ne* *znaj-u*].
I.NOM know-PRS.3SG **that-ACC.SG.N** **that** nothing.SG.GEN NEG know-PRS.1SG
'I know that I know nothing.'

(5) *Ja* *dumaj-u* (****t-o***) [***čto*** *ničego* *ne* *znaj-u*].
I.NOM know-PRS.3SG **that-ACC.SG.N** **that** nothing.SG.GEN NEG know-PRS.1SG
'I think that I know nothing.'

Accordngly, structures with correlatives are possible with emotion predicates like *radovat'* 'please' that are always factive.

We agree with the anonymous reviewer, who argues (referring to Comrie 1971) that not all structures with *to, čto* are necessarily factive (e.g., the verb *somnevat'sja* 'doubt' can host a *to, čto*-clause). Note, however, that it is plausible to set apart structures with *to, čto* and a preposition (the construction with *somnevat'sja* belongs here) and structures without a preposition. It seems that in prepositional structures the use of *to* is often obligatory and therefore the factivity condition is not applicable there. This condition is applicable mainly in structures without a preposition.

On the other hand, according to Letučij (2012) and Knjazev (2009), in some cases structures without the correlative *to* are the only possible variant, because of the semantic role of the complement clause. For instance, only the variant with the correlative can convey the semantics of reason and/or causer, and for stimulus complement clauses, this same variant is the default one; see (6) with a clause of reason.

(6) **Rabot-u* *zamedlja-et* [***čto*** *u* *nego* *net* *slovarj-a*].
job-SG.ACC decelerate-PRS.3SG **that** at he.GEN no dictionary-GEN
'His work is slowing down, as he has no dictionary.' (reason)

The discussion is structured as follows. For all languages under analysis, first, a general overview of the complementizer system is provided. In what follows, complementation strategies are described, beginning with the basic complementizers like *čto* in Russian, continuing with 'process' complementizers like *kak*, and then with indirect question markers (*li* and *dali*), complement uses of adverbial clause markers, irreal complementizers and so on. Apart from that, some more general issues are considered: for instance, the relative order of the matrix and the embedded clause, the status of constructions without overt complementizer, stylictic restrictions and semantic contrasts in the complementizer systems.

2 General facts about Slavic languages

The Slavic languages are divided in three subgroups: East Slavic (Russian, Belorussian, and Ukranian), West Slavic (Polish, Czech, Slovak) and South Slavonic (Bulgarian, Macedonian, Serbian, Croatian, Slovene). The Slavic languages possess a rich case system, except for Bulgarian and Macedonian where nouns, contrary to pronouns, do not have cases at all, when not bearing an article. The Slavic languages are different in terms of word order, although in general they tend to have free word order in declarative clauses. Finally, in all subgroups interrogative pronouns serve as a source of subordinators, such as relative pronouns, complementizers and adverbializers.

3 The complementizer system of Russian

The term 'complementizer' (henceforth COMP) goes back to generative frameworks where the complementizer phrase is held to be the highest extended verb projection associated with sentence illocutionary force (= declarative, imperative and interrogative clause, cf. Bresnan 1970, Zimmermann 2009). COMP is treated as a node in the deep structure or as a functional element present in all finite clauses, either spelled out as a lexical element or realised as a phonologically empty category. In some recent formal studies, COMP has been split into a number of sub-heads (e.g. Rizzi 1997 for Italian). According to the formal approach, COMP is linked to specific word order rules, i.e. it is strongly associated with the clause initial or second position in the clause. In traditional Russian grammaticography, the elements we call COMP are usually treated as conjunctions introducing so-called 'explanatory sentences' (*iz'jasni tel'nye predloženija*). This notion, however, does not point to any particular syntactic status of the clause. More-

over, we address here some constructions which do not fall under the traditional notion '*iz'jasnitel'nye predloženija*' (see e.g. the Syntax of the Academy of Sciences: Švedova 1982: 471–473 and Barnetová et al. 1979: 950–960). We are not aware of any studies specifically dedicated to the description of the entire system of complementizers, though many studies of sentential arguments deal with distribution of adverbs. In the following section we will be dealing with *čto* 'that', *čtoby* 'to, in order to', *kak (by)* 'like, as, as if', *li* 'whether', *kogda* 'when', *esli* 'if' and *kak budto (by)* 'as though'.

There is empirical evidence to assume that the Russian clause syntax has a specific word order slot associated with COMP. The Russian COMP system consists of elements many of which are derived from interrogatives or include an interrogative-based unit. Zimmermann (2009: 502) points out that there is a firm derivational link between COMPs, interrogatives and indefinite pronouns; for example, the following elements are derivationally linked to the complementizer *čto*: the interrogative *čto* 'what', the indefinite-specific *čto-to/nečto* 'something' and the indefinite non-specific *čto-nibud*' 'anything'. We have identified four COMPs, all of which are characterized by a far reaching syntactic fixation; i.e. they obligatorily occupy the initial position or the second position in the subordinate clause.

3.1 The complementizer *čto*

The element *čto* is the most frequent Russian complementizer. It is not restricted to any clear subtype of clausal complement structures and thus can be considered to be the canonical COMP. It occurs after a very wide class of verbal predicates, deverbal nouns and so on. *čto* is derived from and formally identical to the WH pronoun *čto* 'what'. The same unit is also used as a relativizer, quasi-synonymous to *kotoryj* 'which'. In some contexts, it is facultative and can be omitted depending on grounding procedures: the state of affairs encoded in the complement clause can be treated as backgrounded or as foregrounded information.

(7) *Ja* *znaj-u* [**Ø** *on* *rabota-et* *na* *zavod-e*].
I know-PRS.1SG **that** he work-PRS.3SG on factory-LOC.SG
'I know (that) he works at the factory.'

(8) *Sluča-et-sja* [***čto*** *rebenok* *ne* *choč-et* *id-ti* *v* *škol-u*].
happen-PRS.3SG-REFL **that** child not want-PRS.3SG go-INF to school-ACC.SG
'It happens that a child doesn't want to go to school.'

The variant without a complementizer has a colloquial flavor. However, it seems that when the complementizer is omitted, the (originally) embedded clause becomes the pragmatically main one for the speaker, and the matrix predicate serves as a parenthetical / modal word (see Serdobol'skaya 2011).

As pointed out in the Grammar of the Academy of Sciences (Švedova 1982: 472), not only verbs but also certain adjectives, nouns (e.g. *mysl', čto* 'the thought that') and predicatives (e.g. *grustno, čto* 'it is sad that') can take complement clauses of this type.

(9) *Izvestn-o,* [***čto*** *on* *rabota-et* *na* *zavod-e*].
known-SG.N **that** he.NOM work-PRS.3SG on factory-LOC.SG
'It is known that he works at the factory.'

If the matrix clause is negated, with some verbs, such as *pomnit'* 'remember' or *dumat'* 'think, consider', the complementizer *čto* can be changed to *čtoby*, consisting of *čto* + *by* (the conditional/unreal marker). Dobrushina (2012) calls this use of *čtoby* 'unreal' (see the discussion after example [24] below for detail). The complementizer *čtoby* originally goes back to a fusion of the element *čto* with the subjunctive (or conditional) morpheme *by* (our glossing reflects this etymology). In contexts like (10) the choice of *čtoby* instead of *čto* reflects the non-factivity or the probable non-reality of the situation in the embedded clause; this switch of the complementizer under negation can be termed 'negative alternation'.

(10) *Ja* *ne* *pomnj-u,* [***čto-by*** *ty* *mne* *govori-l* *ob* *ėt-om*].
I not recall-PRS.1SG **that-COND** you.NOM I.DAT tell-PST about that-LOC.SG
'I don't recall you telling me about that.'

In complement clauses governed by volitional or other verbs containing a deontic component, the fused COMP *čtoby* is used if the subject of the main clause does not control the reference of the subject of the subordinated clause, i.e. if the subjects refer to different entities marked in example (11) with the two indices *x* and *y* (for the notion of 'control' see Landau 2000 and Farkas 1988, and for 'syntactic obviation' see Avrutin & Babyonyshev 1997 and Szucsich 2010):

(11) ***Ja***$_{\mathbf{x}}$ *xoč-u,* [*čto-by* ***ty***$_{\mathbf{y}}$ *by-l-a* *zdes'*].
I want-PRS.1SG that-COND **you.NOM** be-PST-SG.F here
'I wish you were here.'

In the case of referential identity of the subjects usually the infinitive construction is used which according to our definition does not form a complement clause. The non-overt PRO of the infinitive construction receives its referential interpretation from the subject of the matrix clause; see the indices in example (12). Specific contexts involving mainly the first and second person (rarely the third person),

however, do allow for referential identity with the conditional, as example (13) shows:

(12) ***Ja***$_{\mathbf{x}}$ *xoč-u* [**PRO**$_{\mathbf{x}}$ *by-t'* *zdes'*].
I want-PRS.1SG be-INF here
'I wish to be here.'

(13) ***Ja***$_{\mathbf{x}}$ *xoč-u,* [*čto-by* ***ja***$_{\mathbf{x}}$ *by-l* *zdes'*].
I want-PRS.1SG that-COND **I.NOM** be-PST here
'I wish I were here.'

Examples like (13) serve as counterevidence to the claim of Avrutin and Babyonyshev (1997: 259) that same subject constructions with *čtoby* are in general ungrammatical. The two constructions in (12) and (13) seem to differ in relation to reality: while in (12) the state of affairs is construed as non-real, (13) refers to a state of affairs which does not exist in the time of the utterance, but has the potential to become true. Here belong also cases when the speaker intentionally does not represent himself/herself as a controller of the event:

(14) ***Ja*** *xoč-u* [*čto-by* ***ja*** *okaza-l-sja* *nevinovn-ym*].
I want-PRS.1SG that-COND **I.NOM** find.oneself-PST-REFL innocent-INS.SG
'I wish I would prove innocent.'

(15) ***Ja*** *xoč-u* [*čto-by* ***ja*** *by-l* *bessmertn-ym*].
I want-PRS.1SG that-COND **I.NOM** be-PST immortal-INS.SG
'I wish I would be immortal.'

With object control verbs, the distribution of *čtoby*-clauses vs. infinitive clauses is even more complicated. For instance, both (16) and (17) are possible; in both cases the indirect object of the matrix clauses *nam* controls the reference of the subject of the subordinate clause: non-overt PRO in example (16) and *my* in example (17).

(16) *On*$_{x}$ *vele-l* *nam*$_{y}$ [PRO$_{y}$ ***prij-ti*** *v* *10* *čas-ov*].
he.NOM order-PST us.DAT **come-INF** at 10 hour-GEN.PL
'He ordered us to come at 10 o'clock.'

(17) *On*$_{x}$ *vele-l* *nam*$_{y}$ [***čto-by*** *my*$_{y}$ ***priš-l-i*]**. *v* *10* *čas-ov*].
he.NOM order-PST us.DAT **that-COND** we.NOM **come-PST-PL** at 10 hour-GEN.PL
lit. 'He ordered us that we should come at 10 o'clock.'

(18) *On*$_{x}$ *predloži-l* *nam*$_{y}$ [PRO$_{x+y}$ ***poj-ti*** *v* *kino*].
he.NOM offer-PST us.DAT **go-INF** at cinema
'He suggested going to cinema.'

(19) *On*$_x$ *predloži-l,* *nam*$_y$ [***čto-by*** *my*$_y$ ***poš-l-i*** *v* *kino*].
he.NOM offer-PST us.DAT **that-COND** we.NOM **go-PST-PL** at cinema
'He suggested we go to the cinema.'

These data show that the Russian complementation strategies closely interact with the syntactic devices of argument sharing, referential control and obviation.

Čtoby is also used as a purpose marker in adverbial clauses. The difference is that in these uses, *čtoby* can host either a 'past tense' form or an infinitive. In argument clauses, the combination *čtoby* + infinitive is mainly unavailable (as said above, infinitive without *čtoby* is used to mark control relations). However, in some contexts the combination *čtoby* + infinitive seems to mark arguments rather than adjuncts:

(20) *My* *dogovori-l-i-s'* [***čto-by*** *pojas-ami* ***torgova-t'***].
we.NOM stipulate-PST-PL-REFL **that-COND** belt-INS.PL deal-INF
'We agreed to deal in belts.'

(21) *Kak* *sdela-t'* [***čto-by*** ***uzna-t'*** *nomer* *mobil'n-ogo*]*?*
how do-INF **that-COND** **recognise-INF** number.ACC mobile.phone-GEN.SG
'What do you have to do to get a mobile number?'

In other words, the possibility to form control constructions with the infinitive seems to be a specific feature of the Russian constructions with *čtoby*, which is not fully predictable from the occurrence of *čtoby* as a complementizer.

It is interesting to note that the use of the conditional *by* in the complement clause can trigger semantic coercion in the meaning of the matrix verb: a verb of saying receives the additional deontic component of a demand or request directed towards the referent encoded as dative object. In other words, verbs like *skazat'* 'say' are interpreted as manipulative predicates. This is an interesting case of semantic coercion from the complement to the matrix clause; cf.:

(22) *Ja* *by* *daže* ***skaza-l-a,*** [***čto*** *knig-a* *dlja* *vs-ej* *sem'-i*].
I COND even **say-PST-SG.F** **that** book-SG.NOM for whole-GEN.SG.F family-GEN.SG
'I would even say that the book is for the whole family.' (utterance predicate)

(23) *On* ***skaza-l*** *nam,* [***čto-by*** *my* *podava-l-i* *v* *sud* *na* *kommunal'ščik-ov*].
he.NOM **say-PST** we.DAT **that-COND** we.NOM bring-PST-PL to trial on kommunalshchik-ACC.PL
'He told us to sue the public services.' (manipulative predicate)

The semantic coercion involves in many cases object control, i.e. the indirect object of the matrix verb imposes the reference of the subject of the complement clause. However, there are equally examples where *čtoby* is used with *skazat'* 'say' with the meaning of order or advice, though no co-reference is observed. In (24) we observe an implicit reference to a covert source responsible for bringing about the situation in the sense of Farkas (1988).

(24) *Ja*$_x$ *skaza-l emu*$_y$ [*čto-by bol'še tak-ogo*$_z$ *ne by-l-o*].
I.NOM tell-PST he.DAT that-COND more such-GEN.N not be-PST-SG.N

Example (24) receives the interpretation that 'I asked him (*y*) to bring about the fact that something like that doesn't happen again'. It seems that co-reference between the addressee argument of the main clause and the subject of the embedded clause in examples like (23) results not from strict syntactic rules, but rather from the issues of volitionality and the capability of the agent to influence the state of affairs. *Skazat'* 'say' in the meaning 'order' requires that the event in the embedded clause is controlled by the addressee of the main clause, which is usually the case when the addressee is the subject of the embedded clause.

A more detailed analysis of the marker is proposed by Dobrushina (2012). She demonstrates convincingly that the distribution of *čto* and *čtoby* is far from being trivial. As we have seen, it is usually assumed that *čtoby* is reserved for unreal situations. However, this cannot explain why, for instance, the verb *nadejat'sja* 'hope' does not take *čtoby*, but only *čto*. Dobrushina distinguishes between two types of uses of *čtoby*: one which she calls purpose-like *čtoby* which is characteristic for verbs of desire, necessity and possibility, and the other one called unreal *čtoby*. The first type was exemplified by (13) expressing a wish and by (23) expressing an order. The second type of *čtoby* usually occurs with epistemic verbs, after sentential negation, as in (10) (*ne pomnju, čtoby* 'I don't remember that', lit. 'I don't remember (him) to') or after verbs containing an implicit negative component (*somnevajus', čtoby* 'I doubt if'). The use of the complementizer *čtoby* in (10) is motivated by the sentential negation which implies that the situation in the embedded clause has not necessarily taken place.

For some predicates, this semantic analysis of *čtoby* is not unproblematic. For instance, it can be used with the verb *privyknut'* 'get used to':

(25) *Ja ne privyk* [***čto-by*** *ko mne*
I.NOM NEG get.used.PST.SG.M **that-COND** towards I.DAT
obraščа-l-i-s' po imeni-otčestv-u].
address-PST-PL-REFL by name-paternal.name-DAT
'I am not used to be addressed by my paternal name.'

With this verb, *čtoby* is mainly characteristic for negative contexts. In positive contexts, the verb hosts *čto*-clauses in 436 examples and *čtoby*-clauses in 57 examples excerpted from the Russian National Corpus. In contrast, when *privyknut'* is used with negation, *čtoby* with 24 examples significantly outranks *čto* which is found in only 10 examples. This makes *privyknut'* distributionally similar to verbs like *pomnit'* and *somnevat'sja*. However, the differences are that a) *čtoby* can be used with *privyknut'* even without negation and b) the verb *privyknut'* expresses in some contexts the experiencer's will that the situation become true (i.e. conveys positive attitude toward the situation).

The situation with the verb *ljubit'* 'like; love (it)' is almost converse. This verb can be used with the complementizers *čtoby* and *kogda* 'when'. *Kogda* is found in 705 cases under negation and 588 cases without negation. In contrast, *čtoby* is used in 174 examples under negation and 407 examples without negation. Thus, with this verb, *čtoby* is mainly used in positive constructions. These facts show the semantics of the complement taking predicate influence the polarity sensitivity of *čtoby*.

Interestingly, the epistemic *čtoby* is highly dependent on grammatical person. The tables below show the distribution of *čto* and *čtoby* in first and third person singular contexts with different propositional attitude predicates. It is evident that *čtoby* co-occurs much more frequently with first person than with third person subjects of the complement taking verb.

Table 1: *somnevat'sja* 'doubt'

	čtoby	***čto***
1SG PRS	135	775
3SG PRS	8	211

2-tail Fisher's test: *p*-value < 0.0001

Table 2: *pomnit'* 'remember'

	čtoby	***čto***
1SG PRS	376	380
3SG PRS	16	71

2-Tail Fisher's test: *p*-value < 0.0001

Table 3: *dumat'* 'think'

	čtoby	*čto*
1SG PRS	587	1081
3SG PRS	10	113

2-tail Fisher's test: *p*-value < 0.0001

Table 4: *verit'* 'believe'

	čtoby	*čto*
1SG PRS	72	307
3SG PRS	11	137

2-tail Fisher's test: *p*-value = 0.0008

For all four verbs, the distinction between the first and the third person is statistically significant. The reason for this covariance with person seems to be the following. Using *čtoby* the speaker marks the situation as highly improbable. Correspondingly, the speaker can more easily estimate the probability of the event if s/he coincides with the psychological subject of the complement taking predicate – in other words, when s/he is the one who 'believes' or 'thinks'.

Another class of problematic cases is the verbs of hoping, which do not host *čtoby*, although the event described in complement constructions with these verbs is unreal. The key feature distinguishing 'hope'-verbs which do not host *čtoby* and 'want'-verbs which host *čtoby* seems to be the temporal location of the embedded event. It seems that 'hope'-verbs do not take *čtoby*-clauses because the embedded event can be situated in the past with respect to the mental act; cf. (26). This is, of course, impossible for 'want'-verbs. In (27), only the first reading is accessible ('I want that he bought a ticket [in future]'); the second one ('I want that he has already bought a ticket') is impossible.

(26) *Nadej-u-s'* [*čto* *on* *kupi-l* *bilet*].
hope-PRS.1SG-REFL that he.NOM buy-PST.SG.M ticket.SG.ACC
'I hope that he has already bought a ticket.'

(27) *Ja* *xoču* [*čto-by* *on* *kupil* *bilet*].
I.NOM want-PRS.1SG that-COND he.NOM buy-PST.SG.M ticket.SG.ACC
'I want that he bought a ticket (in future).'
*'I want that he has already bought a ticket.'

Discussing the difference between *čto* and *čtoby*, Brecht (1977) has claimed that the properties of *čtoby* cannot be reduced to those of *čto* + *by*. The class of predicates taking *čtoby* intersects with those taking *čto* (sometimes in the combination with the conditional *by* elsewhere in the complement clause), but does not coincide with it. For instance, verbs like *chotet'* 'want' or *prikazat'* 'order' cannot take *čto* (with or without *by*), but can take *čtoby* (see Khomitsevich 2007).

Further evidence for the perplex distribution of *čtoby* is provided by constructions with the verb *načat'* 'start'. These constructions can host either *čto* or *čtoby*. However, with *čto,* the verb form in the embedded clause usually is identical to that in the matrix clause. With *čtoby*, the verb in the embedded clause is usually in the infinitive:

(28) *Ja* *nača-l* *by* *s* *togo* [***čto*** ***objavi-l*** ***by*** ...]
I.NOM begin-PST COND with that **that** **declare-PST** **COND**
'I would start by declaring ...'

(29) *Ja* *nača-l* *by* *s* *togo* [***čto-by*** ***objavit'*** ...]
I.NOM begin-PST COND with that **that-COND** **declare-INF**
'I would start by declaring ...'

3.2 The complementizer *kak*

The second complementizer obligatorily occupying the first slot in the clause is derived from the interrogative element *kak* 'how'. This element is restricted to verbs of perception and occurs in complements which express an 'object of immediate perception':

(30) *Ja* *vide-l* [***kak*** *mam-a* *celova-l-a* *Santa Klaus-a*].
I see-PST **how** mum-SG.NOM kiss-PST-SG.F Santa Claus-SG.ACC
'I saw Mum kissing Santa Clause.'

In contrast, objects of indirect perception, for instance in reported situations like 'I heard that he had left', are always marked by *čto*:

(31) *Ja* *slyša-l* [***čto*** *on* *uexa-l*].
I.NOM hear-PST.SG.M **that** he.NOM leave-PST
'I heard (I was told) that he had left.'

The opposition of *kak* vs. *čto* seems to follow the pattern proposed by Boye (Boye 2010a, 2010b, 2012: 192): *kak* denotes a state of affairs which is not truth-valued, while *čto* introduces a proposition which is truth-valued. However, the variant with *čto* is the default one and can be also used with complements of immediate

perception. The key feature seems to be that for *kak*, the speaker perceived the situation going on, while for *čto*, it is possible that s/he only observed some parts of it.

As noted by Noonan (2007: 131) and Barnetová et al. (1979: 958), the unit *kak by* has different semantic features. It coincides with the conditional mood and negation in complements of predicates of fearing and worrying, such as *volnovat'sja* 'worry', *bespokoit'sja* 'worry', *bojat'sja* 'fear':

(32) *Boj-u-s'*, [***kak*** ***by*** *na ėt-o* ***ne*** ***uš-l-i***
fear-PRS.1SG-REFL **how** **COND** on this-SG.ACC **not** **leave-PST-PL**
vs-e *ego* *den'g-i*].
all-PL.NOM his money-PL.NOM
'I am afraid that all his money might go on this.'

Note that *kak* cannot mark sentential arguments of these verbs. In general, after verbs of fearing there is a competition between a) *kak* plus conditional and negation, b) *čto* plus conditional and, c) *čto* plus affirmative indicative, the latter being the most frequent and unmarked construction. These constructions differ in the epistemic stance, i.e. in the degree of epistemic distancing from the proposition expressed in the construction. In the case of *kak* plus conditional and negation, the content of the feared state of affairs is portrayed as only remotely possible, whereas the complement clause with *čto* and the affirmative indicative indicates a very real threat; compare example (32) with (33):

(33) *A* *vy* *ne* *boi-te-s'*, [***čto*** *vas* *sam-ogo*
and you.NOM not fear-PRS.2PL-REFL **that** you.ACC self-ACC.SG
vykinu-t]?
throw.out-IND.PRS.3PL
'And you are not afraid they would throw you out?'

Notably, in the construction with *čto*-clause *bojat'sja* has a special (epistemic) use (it means something like 'consider, suppose'). This reading is not available in construction with an infinitive clause or a nominal object. The construction with *čtoby* is rather rare with the verbs under analysis, and therefore it is difficult to discuss its semantic properties.

Another possible analysis is that *kak by* is an independent complementizer. There are three arguments for this analysis. First, *by* in *kak by* can never move, but always occurs immediately after *kak*. This is not the case in free combination of *by* with other complementizers; in example (34), *by* can occupy different positions with respect to *čto*, but not one immediately after it:

(34) *Boj-u-s'*, [*čto* *on* *by* [**čto* *by* *on*] *ne*
fear-PRS.1SG-REFL COMP he.NOM COND COMP COND he.NOM not
spravi-l-sja].
succeed-PST-REFL
'I am afraid he wouldn't be able to do it.'

In other words, *čto* not only forms with *by* the complex unit *čtoby*. There are cases where *čto* + *by* serve as a free combination of units, for instance, when a verb takes *čto* as a complementizer, and the verb in the embedded clause is in the subjunctive mood.

Second, the group of predicates allowing complements with *kak by* does not even intersect with those which allow complements with *kak*. Third, *kak by*-clauses themselves cannot occupy the position before the matrix clause. This is in principle the main feature which distinguishes argument uses of some polysemous markers (*kak by*, *kogda, esli, budto*) from their adjunct uses, which are compatible with preposed dependent clauses. This shows that *kak* has partly lost its independent lexical meaning of 'manner' in argument clauses introduced by *kak by*.

(35) [**Kak* *by* *on* *ne* *opozda-l*], *ja* *boj-u-s'*.
how COND he.NOM not be.late-PST I.NOM fear-PRS.1SG-REFL
Intended: 'I am afraid that he can be late.'

3.3 The interrogative marker *li*

Verbs of speaking, thinking, knowledge and some others can take a complement clause containing the interrogative marker *li*, an enclitic element which functions as an optional marker of non-embedded yes-no-questions; cf. example (36). When *li* is used in a complement clause, the reality status of the complement state of affairs is left open, as in (37). In contrast to *čto* and *kak*, this question marker cliticizes with the first prosodic word in the clause (usually the finite verb) (cf. Franks & King 2000: 189–190):

(36) *Priezžaj-ut* ***li*** *k* *vam* *moskovsk-ie* *antrepriz-y?*
come-PRS.3PL **whether** to you.DAT Moscow-NOM.PL private.theatre-NOM.PL
'Do the private theatre companies from Moscow visit your place?'

(37) *Ja* *ne* *znaj-u*, [*rabota-et* ***li*** *on* *na* *zavod-e*].
I not know-PRS.1SG work-PRS.3SG **whether** he.NOM on factory-LOC.SG
'I don't know if he works at the factory.'

Note that *li* is used even with verbs which do not host indirect questions in the classical sense. It is possible, for instance, with verbs like *bespokoit'(sja)* 'worry / be worried' (see Padučeva 2010 where such occurrences are treated separately from canonical indirect questions):

(38) *Ego* *bespoko-it,* [*ne* *zave-l-a* *li* *ty* *nov-yj*
he.ACC worry-PRS.3SG not make-PST-SG.F whether you.NOM new-M.SG.ACC
roman].
roman
'He is worried that you might have a new love story.'

In this case, what worries the experiencer is the possibility of the positive answer. Contrary to verbs like *dumat'* 'think' (e.g. 'He is thinking whether he is in Moscow or not'), examples with *bespokoit'sja* 'worry' do not presuppose that the experiencer seeks to get an answer to the question – one of the answers is in the focus of attention. See Padučeva (2010) for a detailed description of indirect questions with verbs like *bespokoit'sja* and their correlation with factivity.

This interrogative marker sounds old-fashioned in direct questions, but is very frequent in indirect ones. However, the same type of questions can be formed without *li* – in this case we deal with direct questions, rather than indirect ones:

(39) *Golubčik-i,* *ja* *ne* *ponimaj-u,* [*vy* *v*
dear-PL.NOM I.NOM not understand-PRS.1SG you.NOM in
svoj-em *um-e[?]*]
own-M.SG.LOC mind-SG.LOC
'Guys, I don't understand if you are sane (lit. in your own mind.)'

These structures can either be regarded as indirect questions (in [39], a question is embedded under the verb *ponimat'* 'understand') or as direct questions – in this case, the verb *ponimat'* does not have an object argument. The marker of disjunction *ili* 'or' behaves very similarly to *li* and can be supposed to serve as a marker of indirect questions. For most verbs, the two parallel constructions – the one with *ili* and the one with *li* – co-exist. Compare (40) with (37) above:

(40) *Ja* *ne* *znaj-u,* [*rabota-et* *on* *na* *zavod-e* ***ili*** *v*
I.NOM not know-PRS.1SG work-PRS.3SG he.NOM on plant-LOC.SG **or** in
magazin-e].
shop-LOC.SG
'I don't know if he works at a factory or at a shop.'

Of course, the parallelism of the structures with *li* and *ili* does not make *ili* an indirect question marker *sensu stricto*. However, the very fact that the presence of *ili* makes *li* optional, while in other indirect common questions it is mandatory (see [41] which is ungrammatical because *li* is absent), cannot be ignored:

(41) **Ja ego sprosi-l on prid-et.*
I.NOM he.ACC ask-PRS.3SG.M he.NOM come-PRS.3SG
Intended: 'I asked him whether he would come.'

Sometimes, the two markers even co-occur:

(42) *Ja ne znaju, [rabota-et* ***li*** *on na zavod-e*
I.NOM not know-PRS.1SG work-PRS.3SG **whether** he.NOM on plant-LOC.SG
ili *v magazin-e*].
or in shop-LOC.SG
'I don't know if he works at the factory or in the shop.'

However, there are some cases where *li* can only restrictedly be replaced with *ili*. Consider the complements of *bespokoit'sja* 'worry'; (43) is grammatically absolutely correct, whereas (44) is only marginally possible:

(43) *Menja bespoko-it, [smog* ***li*** *on ses-t' na poezd*].
I.ACC worry-PRS.3SG can.PST.3SG **whether** he.NOM sit-INF on train
'It worries me if he managed to catch the train.'

(44) ???*Menja bespoko-sit, [smog on ses-t' na poezd* ***ili*** *emu*
I.ACC worry-PRS.3SG can.PST he.NOM sit-INF on train **or** he.DAT
priš-l-o-s' id-ti peškom].
have.to-PST-SG.N-REFL go-INF by.foot
'It worries me if he managed to catch the train or he had to go by foot.'

Examples like (43) and (44) make it evident that *li* and *ili* are not strictly parallel. We can claim that the construction with *ili* is more restricted to complements of predicates with a question component *sensu stricto*. *Ili* in complements means that the choice of answer is important for the speaker (e.g. with *ne znat'* the experiencer does not know which of the alternative answers is the correct one). In contrast, the structure with *li* can equally denote the question itself – what annoys the experiencer in the case of *bespokoit'sja* is the question itself, and not the choice of a particular answer.

The combination of *li* and the negation marker *ne* is also used as an indirect question marker. Moreover, with some verbs, such as verbs of fear and worry, *ne + li* is at least equally frequent as *li*:

(45) *Menja bespokoi-l-o [****ne*** *obman-et* ***li*** *on menja*].
I.ACC worry-PST-SG.N **not** deceive-PRS.3SG **whether** he.NOM I.ACC
'It worried me that he could deceive me.'

For instance, with the verb *bespokoit'sja* 'worry', *ne + li* is found in 32 examples, while *li* without *ne* occurred in 29 examples in the Russian National Corpus. It seems that *ne + li* does not differ from *li* without negation. In both cases what

worries the speaker is the question whether the person talked about can deceive him. In contrast, what distinguishes the two constructions seems to be the modality: *ne + li* has more perceivable modal flavor than *li*. While (45) is normal, denoting something like 'It worried me that he *could* deceive me', (46) is grammatical but sounds strange as a question with two symmetrical alternatives; as if the speaker was worried by the question itself whether the person will deceive him or not, with no hint as to the possible scenario, which makes him worry.

(46) *Menja* *bespokoi-l-o* [*obman-et* ***li*** *on* *menja*].
I.ACC worry-PST-SG.N deceive-PRS.3SG **whether** he.NOM I.ACC
'It worried me that he could deceive me.'

3.4 Temporal and conditional subordinators as complementizers

A less obvious case of complementizer is represented by the temporal subordinator *kogda* and the conditional subordinator *esli*[3]. In their principal uses, they are used to form adjunct (adverbial) clauses:

(47) *Kogda* *my* *poš-l-i* *gulja-t'*, *nača-l-sja* *dožd'*.
When we.NOM go-PST-PL walk-INF begin-PST-REFL rain
'When we went for a walk, it began to rain.'

(48) *Esli* *pried-et* *Katj-a,* *ja* *bud-u* *rad*.
if come-PRS.3SG Katia-NOM I.NOM be.FUT-1SG glad.SG.M
'I will be glad if Katia comes.'

However, the same markers also have complementizer uses. They belong to the type described, among others, by Serdobol'skaya (2011) and Letuchiy (2014). In this type of complement construction, the event in the subordinate clause is not just described but is also modified by some aspectual or modal meanings.

(49) *Ne* *ljublj-u,* [***kogda*** *mne* *grubj-at*].
not love-PRS.1SG **when** I.DAT be.rude-PRS.3PL
'I don't like it when people are rude to me.'

(50) *A* *tebe* *ponrav-it-sja,* [***esli*** *s* *toboj* *tak* *postupj-at*].
and you.DAT like-PRS.3SG-REFL **if** with you.INS so behave-PRS.3PL
'And you, will you like it if people behave this way with you?'

3 Diachronically, the subordinator *esli* includes the third person singular present tense form *est'* of the copula verb *byt'* 'be' and the polar question marker *li* (see Section 3.3 above for details).

Kogda expresses an iterative meaning; i.e. the event in the embedded clause is non-episodic and occurs multiple times. Serdobol'skaya (2012) describes this meaning as indefiniteness of the embedded event. *Esli* introduces a modal component to the semantics of the sentence: in (50), as well as in the proper conditional use, *esli s toboj tak postupjat* conveys that the situation described has not come to be true – it is only regarded as a condition of occurrence of the other situation ('In the case if people behave with you in this way, will you like it?'). Besides, *esli*, just as *kogda*, can also mark indefiniteness (multiple occurrences) of the event in the embedded clause.

The relation of *esli* with factivity in contexts like (50) is of special interest. It may seem that *esli* introduces a non-factive predication. However, this appears to be false (see Letuchiy 2014 for detail). The verb *ponravit'sja* 'please' is always factive in the sense that if something (e.g., 'that John came on time') pleases the experiencer, it is true that the event itself took place (John really came on time). What *esli* does is that it resolves a contradiction created by inserting a factive predicate into a non-factive (modal) context: the verb *ponravit'sja* itself presupposes that the embedded event is true, but the context points to the fact that the whole situation has not come to be true. *Esli* can be described as a marker introducing a possible world, and inside this world the predicate retains its factivity.

The difference between these two markers and the other complementizers is that *kogda* and *esli* are compatible with a very small group of verbs. For instance, *kogda*-clauses can only be complements of certain verbs of mental activity and emotional attitude: *ljubit'* 'like, love'; *nravit'sja* 'like'; *nenavidet'* 'hate'; sometimes *ponimat'* 'understand'. The same is true for *esli* which is mainly restricted to the same classes: *nravit'sja* 'like', *udivljat'* 'surprise', *predpočitat'* 'prefer', and so on.

A question that arises is whether complement-like uses of *kogda* and *esli* are always clearly distinguishable from examples where they are used as adverbial subordinators. The answer to this question is negative. For instance, in cases like (51), no clear-cut distinction can be made:

(51) *Mne* *grustn-o,* [***kogda*** *ja* *viž-u* *tako-e*].
I.DAT sad-N.SG **when** I.NOM see-PRS.1SG such-N.SG.ACC
'I feel sad when I see something like this.'

On the one hand, (51) seems to exemplify a complementizer use of *kogda* ('The fact that I see something like this makes me feel sad', where 'I see something like this' is an argument of 'I feel sad'). On the other hand, the sentence can be interpreted as containing adverbial subordination ('In the moments when I see something like this I feel sad'). The only reliable criterion which allows distinguishing between adverbial and complementation uses is based on the order of

clauses. Virtually no restrictions exist in constructions with adverbial subordination, while in examples with *kogda* and *esli* used as complementizers, a strict rule applies. The embedded clause always follows the main one, and never precedes it. Thus, the transformation of (52), where the *kogda*-clause is a sentential adjunct, into a sentence like (53) is possible, while the transformation of (54), where the *kogda*-clause is a complement clause into (55) is not possible.

(52) *Dožd'* *nača-l-sja,* [*kogda* *my* *poš-l-i* *gulja-t'*].
rain begin-PST-REFL when we.NOM go-PST-PL walk-INF
'It began to rain when we went for a walk.'

(53) [*Kogda* *my* *poš-l-i* *gulja-t'*], *nača-l-sja* *dožd'*.
when we.NOM go-PST-PL walk-INF begin-PST-REFL rain.NOM.SG
'When we went for a walk, it began to rain.'

(54) *Ne* *ljublj-u,* [*kogda* *mne* *grubj-at*].
not love-PRS.1SG when I.DAT be.rude-PRS.3PL
'I don't like it when people are rude to me.'

(55) [**Kogda* *mne* *grubj-at*], *ja* *ne* *ljublj-u.*
when I.DAT be.rude-PRS.3PL I.NOM not love-PRS.1SG
Intended: 'I don't like it when people are rude to me.'

Adverbial markers with complementizer functions have one property that distinguishes them from all other complementizers. They do not have a variant with the correlative *to*. The variant *to esli* 'that + if' does not occur at all in the corpus, while the variant *to kogda* can only occur in case *kogda* introduces an indirect question.

3.5 The elements *budto (by)* and *kak budto*: marginal complementizers or a case of complementizer omission?

Barnetová et al. (1979: 954–955) mention that the element *budto (by)* 'as if' and the nearly analogous marker *kak budto (by)* 'as though' can be used as complementizers. These markers usually introduce an adjunct in a comparative construction where one situation is compared with another, as in (56). This construction has given rise to an evidential particle with inferential meaning, as in (57) (cf. Letučij 2008: 225).

(56) *Ona* *zasmeja-l-a-s'* [***kak*** ***budto*** *prozvene-l* *kolokol'čik*].
she laugh-PST-F.SG-REFL **how** **as.if** chime-PST bell
'She burst out in laughter, like the chime of a bell.'

(57) *Tut* ***budto*** *by* *vs-e* *normal'n-o*.
here **as.if** COND everything-NOM.SG normal-SG.N
'Everything seems to be fine here.'

In certain contexts, after verbs of saying or mental activity like *sni'tsja* 'appear in a dream', *dumat'* 'think' or *kazat'sja* 'seem', *budto (by)* can appear in the initial position of a complement clause. In this case, the speaker indicates that the proposition encoded in the subordinate clause is based on inference or on what other people have said (reportative). For *kak budto*, uses like this are less typical, but also found by a Google search in Internet.

At the same time, *budto (by)* and *kak budto* express epistemic distancing from the communicated propositional contents (more details in Letučij 2008 and Wiemer 2008: 347–353); see (58)–(60). In example (60), for instance, the speaker apparently knows that the proposition is not true (in reality the city is not under the water).

(58) *Agassi* ***utveržda-et,*** [***budto*** ***by*** *ponjati-ja* *ne* *ime-et*].
Agassi **claim-PRS.3SG** **as.if** **COND** concept-GEN.SG not have-PRS.3SG
kakov-y *ego* *rejtingov-ye* *šans-y*.
which-NOM.PL his rating-NOM.PL chance-NOM.PL
'Agassi claims that (lit. 'as if') he didn't have a clue about his rating chances.'

(59) *Zrja* *tol'ko* *ona* ***duma-et,*** [***budto*** *ej* *strašnee*].
for.nothing only she **think-PRS.3SG** **COMP** she.DAT scarier
'For nothing she thinks as if she were more scared.'

(60) ***Pokaza-l-o-s',*** [***kak*** ***budto*** *ves'* *gorod* *uše-l* *pod* *vod-u*].
seem-PST-SG.N-REFL **as** **though** all city go-PST under water-ACC.SG
'It seemed that the whole city sank (lit. went under water).'

According to Letučij (2008: 222–224), there seem to be very subtle semantic contrasts between *budto* and its combination with the conditional particle *budto by*. Whereas the former is not only used after verbs of saying, but also after verbs of mental activity, like *dumat'* 'think' or *snit'sja* 'dream', the latter is restricted to verbs of saying. As a result, only *budto by* renders quotative meaning. Letučij (2008: 229–230) points out that the usage of *budto (by)* in these contexts allow for two different syntactic analyses; he himself treats it as an evidential conjunction that would be equivalent to a COMP. There is some evidence to assume, however, that we are dealing with an evidential particle and a case of COMP omission: we find co-occurrences of the unmarked COMP *čto* plus *budto*; see example (61).

(61) *Kaza-l-o-s'*, **[*čto* *budto*** *vide-l-i-s'* *včera*].
seem-PST-N-REFL **that as.if** see-PST-PL-REFL yesterday
'It seemed as if they had met the day before.'

Another circumstance speaking against treating *budto* as a COMP is the far-reaching synonymy with the evidential particle: we would not like to treat two elements with the same function and the same word order position as homonyms.

On the other hand, the evidence for treating *budto* as a COMP is even stronger. If we treat *budto*-constructions as a case of complementizer omission, we should take them to occur in sentences without a complementizer. As we saw in Section 1, the (arguably omitted) complementizer *čto* is compatible with the correlative pronoun *to*; the *budto*-units discussed here are, on the other hand, incompatible with this correlative element. This, however, can be due to the irrealis semantics of *budto* which makes the contexts of their use different from those where *to* occurs. Yet, there are two crucial properties, which distnguish *budto*-structures from structures without complementizers.

1. Modification of nouns

In general, structures without complementizers can attach to verbs but are very marginal with nouns. In contrast, *budto*, as well as other complementizers, can depend on nouns; cf. the following examples with *čto* 'that':

(62) *On* ***ponima-l,*** **[*čto*** *proigra-l*].
he.NOM **understand-PST COMP** lose-PST.SG.M
'He understood that he had lost.'

(63) ***Ponimani-e,*** **[*čto*** *on* *proigra-l*] *rasstraiva-l-o* *ego.*
understanding-NOM.SG that he.NOM lose-PST.SG.M upset-PST-SG.N he.ACC
'The feeling that he had lost upset him.'

(64) *On* ***ponima-l,*** [*Vasj-a* *ne* *prid-et*].
he.NOM **understand-PST** Vasja-NOM.SG not come-PRS.3SG
'He understood (that) Vasja wouldn't come.'

(65) ****Ponimani-e,*** [*Vasj-a* *ne* *prid-et*], *serdi-l-o* *ego.*
understanding-NOM.SG Vasja-NOM.SG not come-PRS.3SG annoy-PST-SG.N he.ACC
Intended meaning: 'The feeling (that) Vasja wouldn't come annoyed him'

Sentence (63) with a verbal noun is as good as the version with the finite verb in (62). The situation changes in the next examples, where no overt complementizer is present. The structure with the finite verb in (64) sounds well, but the nomi-

nalized variant in (65) is ungrammatical or can be used only under very special intonational conditions.[4]

The *budto*-contstruction behaves like structures with *čto* and other complementizers.

(66) *On* *čuvstvova-l,* **[*budto*** *ego* *obmanu-l-i*].
he.NOM feel-PST **as.if** he.ACC deceive-PST-PL
'He felt that he was deceived.'

(67) *U* *menja* *vsj-u* *žizn'* ***čuvstv-o,*** **[*budto*** *ja* *sebja* *pered*
at I.GEN all-F.SG.ACC life **feeling-SG.NOM** **as.if** I.NOM self.ACC before
nim *ne* *reabilitirova-l*].
he.INS not rehabilitate-PST
'All my life, I have a feeling that I haven't rehabilitate myself before him'

Thus, in this respect *budto* behaves similarly to *čto* and other complementizers. The possibility of examples like (67) shows that *budto* is a complementizer, since the structure with *budto* behaves like canonical constructions with complementizers, such as *čto* and *kak*.

2. Insertion into the main clause

In general, Russian biclausal structures with clausal complements containing no complementizers do not admit insertion of the complement clause into the main clause: the construction in (72) becomes ungrammatical when the embedded clause is inserted in the main clause, as in (73). In this they are similar to coordinate structures that do not allow insertion either (see Testelec 2001 for details). Example (68) becomes ungrammatical when one of the coordinated clauses is inserted in the other; see (69). Sentences with adverbial subordinators, on the other hand, admit this type of insertion; cf. (70) and (71) which are both possible.

(68) *Vasj-a* *pozvoni-l* *Pet-e,* [*i* *oni* *obsudi-l-i* *ėtot* *vopros*]
Vasja-NOM call-PST Petja-DAT and they discuss-PST-PL this.ACC.SG question
'Vasja called Petja, and they discussed this issue.' (coordinate structure)

(69) **Vasj-a,* [*i* *oni* *obsudi-l-i* *ėtot* *vopros*], *pozvoni-l* *Pet-e.*
Vasja-NOM and they discuss-PST-PL this.ACC.SG question call-PST Petja-DAT
Intended: 'Vasja, and they discuss this question, called Petja.'

4 In fact, event nominalizations are usually taken to reflect the properties of the corresponding verbal construction (see Baker & Vinokurova 2009), but there is some evidence showing that some types of arguments are blocked for deverbal nouns (see Letučij 2012).

(70) *Vasj-a* *pozvoni-l* *Pet-e,* [*čto-by* *obsudi-t'* *ėt-ot*
Vasja-NOM call-PST Petja-DAT COMP-COND discuss-INF this-ACC.SG
vopros].
question
'Vasja called Petja to discuss this issue.' (sentence with an adverbial clause)

(71) *Vasj-a,* [*čto-by* *obsudi-t'* *ėt-ot* *vopros*], *pozvoni-l* *Pet-e.*
Vasja-NOM COMP-COND discuss-INF this-ACC.SG question call-PST Petja-DAT
'To discuss this question, Vasja called Petja.'

(72) *Vasj-a* *mne* *skaza-l,* [*on* *ne* *pried-et*].
Vasja-NOM I.DAT tell-PST he.NOM not come-PRS.3SG
'Vasja told me (that) he wouldn't come.' (clausal complement without complementizer).

(73) **Vasj-a* *skaza-l,* [*on* *ne* *pried-et*], *tol'ko* *tebe.*
Vasja-NOM tell-PST he.NOM not come-PRS.3SG only you.DAT
Intended meaning: 'Vasja told only to you (that) he wouldn't come'

In this respect, clauses with *budto* resemble structures with complementizers. The embedded clause can be inserted in the main one. The resulting sentences often sound awkward, though not ungrammatical:

(74) *Vasj-a* *govori-l* *mne,* [*budto* *Pet'-a* *uexa-l*].
Vasja-NOM tell-PST I.DAT as.if Petja-NOM leave-PST
'Vasja told me that Petja had left.'

(75) *Vasj-a* *skaza-l,* [*budto* *Petj-a* *uexa-l*], *tol'ko* *tebe*]
Vasja-NOM tell-PST as.if Petja-NOM leave-PST only you.DAT
'Vasja told only you that Petja had left.'

The marker *budto* and, less frequently, *budto by* can participate in the 'negative alternation'. In other words, they sometimes occur under negation instead of the canonical complementizer *čto* and its derivate *čtoby* (recall the comments on example [10] in Section 3.1). Some verbs of logical connection, such as *značit'* 'mean' and *dokazyvat'* 'prove', can host *budto*-clauses when negated, but only *čto*-clauses without negation:

(76) *No* *ėt-o* *ne* *znač-it* [*budto* *Vasilij* *ne* *sposoben* *greši-t'*].
but this-NOM.SG.N NEG mean-PRS.3SG as.if Vasilij.NOM NEG able.SG.M sin-INF.
'But this does not mean that Vasilij is unable to sin.'

(77) **Ėt-o* *znač-it* [*budto* *Vasilij* *ne* *sposoben* *greši-t'*].
this-NOM.SG.N mean-PRS.3SG as.if Vasilij.NOM NEG able.SG.M sin-INF.
Intended: 'It means that Vasilij is unable to sin.'

As opposed to *budto*, the two other units which are also often found in the beginning of complement clauses – *vdrug* which originally means 'unexpectedly' and *vrode* (and its variants *vrode by* and *vrode kak*) 'similar, like', do not behave like complementizers. Clauses with them cannot be inserted in the main clause or modify nouns:

(78) *Ja* *za* *nee* *boja-l-sja,* [***vdrug*** *zametj-at*].
I.NOM for she.ACC fear-PST-REFL **unexpectedly** notice-PRS.3PL
'I was afraid for her: what if someone notices her.'

(79) *Mne* *govori-l-i,* [***vrode*** *ty* *samovar* *prodae-š*].
I.DAT say-PST-PL **like** you.NOM samovar sell-PRS.2SG
'I was told that you were selling your samovar.'

(80) **Boja-l-sja* [*vdrug* *zametj-at*] *ne* *Vasj-a* *a*
fear-PST-REFL suddenly notice-PRS.3PL not Vasja-NOM.SG but
Petj-a.
Petja-NOM
Intended: 'It was Petja, and not Vasja, who was afraid that he would be noticed.'

(81) **Strax* [*vdrug* *zametj-at*] *by-l* *nevynosim.*
fear.NOM suddenly notice-PRS.3PL be-PST unbearable
Intended: 'The fear that he will be noticed was intolerable.'

Note that in Bulgarian and Polish the process of re-analysis of such modal and comparative markers as the Russian *budto* is not advanced to the same degree. In Bulgarian, for instance, there are two markers of comparison which bear also flavor of irreality: *sjakaš* 'as if' and *kato če (li)* 'as if', which are not used as complementizers. Constructions with predicates like *struva se* 'seem', *mislja* 'think', *sъnuvam* 'dream' parallel to Russian ones are impossible in Bulgarian (see Section 5); see (82) and (83)

(82) **Sъnuva-še* [*kato če* *li* *e* *car*].
dream-IPFV.3SG as.though whether be.PRS.3SG tsar
'He dreamt that he was the tsar.'

(83) **Misl-i* [*sjakaš* *tova* *ne* *e* *istina*].
think-PRS.3SG as.if this not be.PRS.3SG true
'He thinks that it's not true.'

However, we are grateful to Petar Kehayov for pointing out that the element *vse edno* 'all the same, (as if)' can be used both as a modal comparative marker and a complementizer:

(84) *Šte* *mi* *ob'asn'ava* [***vse edno*** *če* *provincija-ta* *izxranva-l-a*
FUT I.DAT explain.PRS.3SG **as.if** COMP province-DEF.F.SG feed-EVID.PRS-F.SG
Sofija].
Sofia
'(He) will explain me that (lit. as if) the province is feeding Sofia.'

3.6 WH interrogatives in complementizer position

Apart from the complementizers described above, Russian allows all WH elements to occupy the COMP slot in the clause and function as markers of a complement clause. These elements differ in relation to the syntactic position within the subordinate complement clause; whereas complementizers like *čto* neither fill a syntactic valence slot opened by the finite verb, nor function as adjuncts, interrogative elements like *kto* 'who' can function as arguments of the embedded predicate (for more details see Meyer 2004); cf.:

(85) *Ja* *znaj-u,* [***kto*** *rabota-et* *na* *zavod-e*].
I know-PRS.1SG **who.NOM** work-PRS.3SG on factory-SG.LOC
'I know who works at the factory.'

The function of embedded wh-questions is close to that of yes/no-embedded questions with *li* (see Section 3.3 above). In both types of structures, the embedded question does not render the illocution of a question. The function of the question structure is to leave a semantic variable unspecified which leads to a certain type of indefiniteness.

It was mentioned in Section 1 that interrogative pronouns are also compatible with the correlative pronoun *to*. With some verbs, either variant – with and without the correlative – can be used. For instance, the verb *znat'* 'know' can host a clause without the correlative, as in (85), as well as a clause with the correlative. In contrast, some verbs, such as *ustraivat'* 'be tolerable, be acceptable' and *udivit'* 'surprise' are impossible with interrogative pronouns without the correlative *to*, or the two variants are not semantically equivalent. In (86) for instance, the variant without *to* is impossible or highly colloquial.

(86) *Menja* *ne* *ustraiva-et* ???(*t-o*) [*komu*
I.ACC NEG be.acceptable-PRS.3SG that-NOM.SG.N who.DAT
poruče-n-o *stroitel'stv-o*].
assign-PTCP.PASS.PST-SG.N construction-SG.NOM
'I cannot accept the fact that the construction works were assigned to this person.'

Interrogative pronouns are compatible with verbs of (acquisition of) knowledge: for instance, *sprosit'* 'ask', or *znat'* 'know' provoke or provide new information. In contrast, verbs of emotional attitude, such as *ustraivat'* or *udivit'*, occur with presupposed complements: they mark attitudes toward certain facts. In this sense, the latter are 'more factive' than the former. This is why *to* must be used with verbs of emotional attitude.

3.7 Combinations of complementizers

It is worth pointing out that a combination of two COMP elements seems to never occur in Russian. For instance, the combination of *čto + budto* is possible, but in cases like this, *budto* can be analyzed as an adverbial. In examples like (87), *budto* occupies a position in the middle of the embedded clause, and not immediately following *čto*:

(87) *Govori-l-i* [***čto*** *on* ***budto*** *by* *umer*].
say-PST-PL **that** he.NOM **as.if** SUBJ die.PST.SG.M
'People were saying that (lit. 'as if') he had died.'

For *kak budto*, the situation is a bit more complicated. This unit could in principle be analyzed as a combination of two subordinators, yet, it is mostly used in comparative adjuncts, as in (88), and only rarely in complement clauses, as in (89):

(88) *Ona* *zasmeja-l-a-s'* [***kak*** ***budto*** *prozvene-l-Ø* *kolokol'čik-Ø*].
she laugh.PFV-PST-F.SG-REFL **how** **as.if** chime-PST-SG.M bell-SG.NOM
'She burst out in laughter, like the chime of a bell.'

(89) *Pokaza-l-o-s',* [***kak budto*** *ves'* *gorod-Ø* *uše-l*
seem-PST-N-REFL **as though** all.M.SG.NOM city-NOM.SG go-PST.SG.M
pod *vod-u*].
under water-ACC.SG
'It seemed that the whole city sank (lit. went under water).'

However, the combination occurring in (89) is semantically not compositional and can hardly be regarded as a combination of complementizers.

3.8 Word order

Russian is considered to be a language with 'free word-order' which means that the order of verbal arguments (including clausal complements) with respect to

each other is not rigid. For instance, examples (90) and (91) show that the clausal complement can occur on both sides of a nominal constituent.

(90) *Ja* *uzna-l,* [*čto* *urok-ov* *ne* *bud-et*], *ot*
I. NOM find.out-PST COMP lesson-GEN.PL not be-PRS.3SG from
direktor-a *škol-y.*
director-GEN.SG school-GEN.SG
'I learnt from the school director that there will be no classes.'

(91) *Ja* *uzna-l* *ot* *direktor-a* *škol-y,* [*čto*
I. NOM find.out-PST from director-GEN.SG school-GEN.SG COMP
urok-ov *ne* *bud-et*].
lesson-GEN.PL not be-PRS.3SG
'I learnt from the school director that there will be no classes.'

For all complementizer structures in question, the position immediately after the verb is the default one, but usually two other positions are possible as well: after another argument, compare (92) and (93), and in the absolute beginning of the sentence; compare (95) and (96) (though the latter option can render some sentences awkward, as in [94]):

(92) *On* *uzna-l* *u* *Vas-i,* [*kto* *ukra-l* *den'gi*].
he.NOM ask-PST at Vasja-GEN who.NOM steal-PST money.ACC
'He asked Vasja who stole the money.'

(93) *On* *uzna-l,* [*kto* *ukra-l* *den'gi*], *u* *Vas-i.*
he.NOM ask-PST who.NOM steal-PST money.ACC at Vasja-GEN.SG
'He asked Vasja who stole the money.'

(94) ?[*Kto* *ukra-l* *den'gi*], *on* *uzna-l* *u* *Vas-i.*
who.NOM steal-PST money.ACC he.NOM know-PST at Vasja-GEN.SG
'He asked Vasja who stole the money.'

(95) *On* *uzna-l* *ot* *Vas-i,* [*čto* *vmešiva-t'-sja* *ne* *nado*].
he.NOM know-PST from Vasja-GEN.SG that meddle-INF-REFL not necessary
'He learnt from Vasja that there is no need to meddle in.'

(96) [*Čto* *vmešiva-t'-sja* *ne* *nado*], *on* *uže* *zna-l.*
that meddle-INF-REFL not necessary he already know-PST
'He already knew that there is no need to meddle.'

There are only two restrictions which are regularly observed. First of all, clausal complements occupying an indirect (oblique) object position are rare in the absolute beginning (see [97]), as opposed to clausal complements in direct object position, as in (96) above.

(97) [*Čto Vasj-a priexa-l], Petj-a by-l očen' rad.
that Vasja-SG.NOM come-PST Petja-SG.NOM be-PST very glad
Intended: 'Petja was very glad that Vasja came.'

Second, complements containing *čtoby* are prohibited or very undesirable in the beginning of the sentence (cf. [98] and [100]), contrary to adverbial purpose clauses with *čtoby* which are frequent in this position (cf. [99] and [101]).

(98) *[Čto-by my otdoxnu-l-i], chote-l-i rodstvennik-i.
that-COND we.NOM rest-PST-PL want-PST-PL relatives-PL.NOM
Intended: 'Our relatives wanted us to have a rest.'

(99) [Čto-by otdoxnu-t'], on poecha-l na mor-e.
that-COND rest-INF he.NOM go-PST to sea-SG.ACC
'He went to the sea to have a rest.'

(100) Rodstvennik-i xote-l-i, [čto-by my otdochnu-l-i].
relatives-PL.NOM want-PST-PL that-COND we.NOM rest-PST-PL
'Our relatives wanted us to have a rest.'

(101) On poexa-l na mor-e [čto-by otdoxnu-t'].
he.NOM go-PST on sea-ACC.SG that-COND rest-INF
'He went to the sea to have a rest.'

3.9 Semantic contrasts between complementizers

Summing up, the Russian complementizer system shows evidence for the following semantic contrasts:

- 'object of immediate perception' (*kak*) vs. 'knowledge acquired' (*čto*)
- degrees of epistemic distancing (*čto* vs. *budto, kak budto*)
- epistemological contrast: indefiniteness contrasts (contrasts related to speaker's knowledge: 'speaker knows *p*' [*čto*] vs. 'speaker does not know *p*' [*čtoby*])
- 'assertion, informative utterance' (*čto*) vs. 'utterance of demand' (*čtoby*)

For some complementizers, the semantic contrasts result from more general features. For instance, we can assume that *čtoby* expresses irreality of the embedded situation. This is why this marker is used in the context of demand (the actions which the addressee is demanded to carry out have not yet become true) and negation (the event described in the embedded clause is not necessarily actualized).

For some other complementizers, as *budto* and *kak budto*, the contrast with *čto* in terms of epistemic distancing seems to result from their semantics as comparative markers. By using these markers, the speaker compares two situations ('He was wandering around, as if he was waiting for somebody') implying that the situation described in the embedded clause is not true (the embedded situation only resembles another, true situation).

3.10 The role of complementizer omission in reanalysis

We would like to conclude the section on Russian complementizers with a short digression dedicated to a specific type of language change which is accompanied or even triggered by complementizer omission: the transition of modal verbs into epistemic adverbs. In many languages epistemic markers are composed from two elements: a modal marker (a verb, an adverbial, and so on) and a copula-like predicate or a complementizer, as *da* in Serbian/Croatian. A known example is English *maybe* which is derived from the modal *may* plus the copula verb *be*. Ramat and Ricca (1998), who analysed sentence adverbs in a wide range of European languages, found that epistemic adverbs often emerge either through the fusion of a modal with a complementizer or through conversion of a modal:

- 'modal.3SG' + 'be'; e.g. English *maybe* (← *may + be*) 'perhaps';
- 'modal.3SG' + 'happen'; e.g. Dutch *misschien* 'perhaps';
- 'modal.3SG' + COMP; e.g. Serbian/Croatian *možda* 'perhaps';
- conversion of a modal; e.g. Romanian *poate* 'perhaps' or 'can.PRS.3SG'.

This development found in nearly all European languages can be analysed on the basis of Russian *možet (byt')*, which goes back to the third person singular present tense of the modal verb *moč'* 'can' plus infinitive. In Hansen (2010) we focused on the constructional aspects of this language change with the aim to show the close interaction of semantics on the one, and argument structure on the other hand, in order to shed light on the specific role of complementizer omission. In the first stages, we are dealing with the transition from a modal construction (NP_{Nom} + Modal + VP_{Inf}) into a complement clause governed by a modal (the so-called 'focal evaluative construction'):

(102) *Odin* *protez* *mož-et* [*služi-t'* *opor-oj* *dlja*
one implant.SG.NOM can-PRS.3SG serve-INF support-SG.INS for
tr-ex *zub-ov*].
three-GEN.PL tooth-GEN.PL
'One of those implanted dentures can function as support for three teeth.'

(103) *Mož-et* *by-t',* [***čto*** *ėto* *problem-a* *ne*
can-PRS.3SG be-INF **that** this problem-SG.NOM not
fizičesk-aja, *a* *psichičesk-aja*].
physical-F.SG.NOM but psychological-F.SG.NOM
'It is possible that this is not a physical, but a psychological problem.'

In a later stage, speakers start omitting the complementizer which leads to syntactic ambiguity: the structure can either be interpreted as a matrix predicate governing a complement clause with the ellipsis of the complementizer or as a parenthetical insertion into a main clause.

(104) ***Mož-et*** ***by-t',*** *ėto* *problem-a* *ne* *fizičesk-aja,*
can-PRS.3SG **be-INF** this problem-SG.NOM not physical-F.SG.NOM
a *psichičesk-aja*.
but psychological-F.SG.NOM
'It is possible that this is not a physical, but a psychological problem.'
or
'Maybe, this is not a physical, but a psychological problem.'

For some modal markers, after the omission of complementizers, the copula-like verb can be omitted too. For example, in Russian, the modal verbs *možet* '(he) can' and, in dialectal and non-educated colloquial speech, *dolžno* can be used alone. This resembles modal conversion, as in Romanian (see above):

(105) *Mož-et,* *on* *bole-et?*
can-PRS.3SG he.NOM be.ill-PRS.3SG
'Perhaps, he is ill?

This reanalysis is accompanied by a shift in grounding procedures. The complement clause carries a specific focal evaluative component: it expresses a speaker-based evaluation of the state of affairs encoded in the subordinate clause. The evaluation or perception is treated as foregrounded, and the state of affairs as backgrounded information (cf. Nuyts 2001 for West Germanic). In the adverbial construction, however, the figure-ground relation is reversed and the modal statement in relation to the state of affairs is treated as backgrounded information. Hansen (2010) argues that it is the omission of the COMP which triggers the syntactic reanalysis leading to the rise of a sentence adverb. Note that structures where the verb is a so-called *vvodnoe slovo* (a parenthetical) are different with respect to the degree to which the evaluation or perception predicate is foregrounded. Compare, for instance, (106) and (107):

(106) [*Ty,* *ja* *viž-u,* *ne-ploxo* *razvlek-sja*].
you.NOM I.NOM see-PRS.1SG not-bad divert.oneself.PST-REFL
'I see that you diverted yourself very well!'

(107) *Vіž-u,* [*sto-it* *Andre-j* *Voznesensk-ij*].
see-PRS.1SG stand-PRS.1SG Andrei-NOM.SG Voznesenskij-NOM.SG
'And I see that Andrei Voznesenskij stands there.'

In (106), the perception verb is fully backgrounded. This is clear from the fact that the clause can participate in co-ordinate structures like 'They went to a night club and, as I see, diverted themselves very well'. In contrast, (107) illustrates a sort of partial re-analysis of the structure. Though the complementizer is lost and the verb behaves as a parenthetical, the perception verb is not backgrounded. For instance, a coordinate structure like (108) is possible:

(108) *Zaxož-u* *v* *bar* *i* *vіž-u,* [*sto-it* *Andre-j*
enter-PRS.1SG in bar and see-PRS.1SG stand-PRS.1SG Andrei-NOM.SG
Voznesensk-ij].
Voznesenskij-NOM.SG
'I am entering the bar and I see that Andrei Voznesenskij stands there.'

In (108), *videt'* 'see' behaves as a full-fledged lexical verb and is coordinated with the verb *zachodit'* 'enter'. In contrast, in (106) where the re-analysis went further, and the perception verb is backgrounded, coordinate structure, analogous to (108) cannot be build based on this sentence (for more details of the historical development see Hansen 2010).

4 The complementizer system of Polish

Due to the lack of space, we will restrict our description of the Polish complementizer system to a relatively coarse outline of its main features. Polish has a complementizer system which shares many features with Russian; i.e. the system consists of function words occupying the initial or second position in the clause. Like in Russian, all COMP elements (with one exception) are derived from interrogatives. In the following, we shall briefly present the main uses of the individual elements. We will discuss the canonical complementizers *że(by), aby, jak, czy* and the peripheral complementizer *kiedy*.

4.1 The complementizer *że* and its variants

The main or semantically unspecified complementizer is the element *że*, which in contrast to the Russian *čto*, is not formally identical with a WH pronoun. *Że* developed from the masculine form of the Proto-Slavic interrogative pronoun *jьže*

'which' which is no longer in use in Polish. The new subjunction *iże* emerged after the loss of the reduced vowel *jь*. This subjunction was retained only in the West Slavic languages (Czech/Slovak *že*). In Polish, the form *że* more or less replaced the synonymous forms *iż/iże* and *eż/eże* (Decyk-Zięba & Dubisz 2008: 268). Nowadays, we still find the form *iże*, but only in literary registers. In contrast to the neutral complementizer *że*, *iż(e)* is stylictically marked, but due to the common etymology there are no semantic differences between them. Although the Russian *čto* and the Polish *że* are derived from different lexical sources, we can state that both have evolved from WH pronouns. The Polish *że* occurs not only with verbs, see (108), but also with predicatively used adjectives and nouns, see (110).[5]

(109)	***Wie-m,***	[***że***	*to*	*brzm-i*	*wulgarnie* ...]	
	know-PRS.1SG	**COMP**	this	sound-PRS.3SG	vulgar	
	'I know that this sounds vulgar ...'					

(110)	*Tylko*	***szkoda,***	[***że***	*nie*	*by-ł-o*	*mnie*	*na*	*tym*	*spotkani-u*].
	only	**pity**	**COMP**	not	be-PST-N	I-GEN	on	this	meeting-LOC.SG
	'It's a pity that I haven't been to this meeting.'								

Like in Russian non-factual complement clauses, the complementizer *że* is fused with the clitic conditional morpheme *by*. In contrast to Russian, however, the compound complementizer bears person and number agreement with the subject. Furthermore, in Polish the conditional morpheme *by* can function as a complementizer itself (see example [116] below), and can fuse with a greater choice of complementizers than its Russian counterpart. These fused complementizers are: *żeby, aby, ażeby, iżby*. In the traditional Polish linguistic literature, the type of sentences introduced by these complementizers is called 'final adjunct clause'. As in the case of *że* and *iże*, the difference between these forms is stylistic, not semantic in nature. Therefore, we can treat them as synonyms but owing their stylistic differences they are not always interchangeable. In colloquial Polish, in the journalese and in the administrative language, the use of *żeby* is common. In these cases and also in casual talking, overusing *aby* sounds pretentious (Markowski 2000: 3). Using *aby, ażeby* and *iżby* in colloquial Polish is embarrassing. Prescriptive guidebooks advise the speaker to use these forms only in written texts. Markowski (2000: 298) points out that *iżeby* is "old-fashioned" (*przestarzały*). Stylistically somewhere between *żeby* and the other forms is *by*, although this form is also reserved for written texts. This does not mean

5 The Polish examples are taken from the Internet and from the National Corpus of Polish (*Narodowy Korpus Języka Polskiego*). As in the case of Russian in some cases we have constructed examples in order to illustrate specific functional contrasts.

that speakers use only the form *żeby* in oral communication. If they use the other forms, they are aware of the stylistic difference and they use them in order to achieve specific stylistic effect, e.g. to sound more emphatic, educated or even pretentious. In written texts, these COMP-forms are completely synonymous. Another complementizer used in colloquial speech is *coby*. This form cannot be used in formal speech or writing.

In Polish, the fused (compound) complementizer has to be used if the main clause and complement clause contain different subjects (the so-called 'obviation'); cf. (111) with ungrammatical (112).

(111) *Chc-ę* [***że-by-ś/a-by-ś/a-że-by-ś/iż-by-ś*** *przyszed-ł* *do* *dom-u*].
want-PRS.1SG **COMP-COND-2SG** come-PST to house-GEN.SG
'I want you to come home.'

(112) **Chc-ę,* [***że*** *przyszed-ł=**by-ś*** *do* *dom-u*].
want-PRS.1SG **COMP** come-PST=**COND-2SG** to house-GEN.SG
'I want you to come home'

As in Russian (see examples [13]–[15] above), certain contexts involving first or second person allow for the usage of *żeby* in same subject constructions. Here the speaker represents himself or herself as not being in control of the event, as in:

(113) *Czasami* *chcia-ł-by-m,* [***że-by-m*** *cię* *nigdy* *nie* *spotka-ł*].
sometimes want-PST-COND-1SG. **COMP-COND-1SG** you.ACC never not meet-PST.
'Sometimes I wished I never had met you.'

(114) *ma-m* *laptop-a,* *chcia-ł-by-m,* [***że-by-m*** *móg-ł*
have-PRS.1SG laptop-ACC want-PST-COND-1SG **COMP-COND-1SG** can-PST
gra-ć *w* *inn-e* *gr-y*].
play-INF in other-ACC.PL game-ACC.PL
'Hi, I have a laptop; I wish I could play other games.'

4.2 The complementizer *aby*

Apart from *żeby,* one finds the stylistically more restricted element *aby*, which is a fusion of the conjunction *a* 'and' and the conditional element *by*:

(115) *Chc-ą* *bowiem,* [***a-by*** *ich* *dzieck-o* *spotka-ł-o*
want-PRS.3PL since **and-COND** their child-NOM.SG meet-PST-N
się *z* *chrześcijańsk-ą* *kultur-ą*].
REFL with Christian-INS.SG culture-INS.SG
'They thus want their children to be in contact with Christian culture.'

It is interesting to note that in certain contexts the Polish form *by*, which coincides with the conditional clitic can be used as a complementizer; see (116). As we know from the discussion in Section 3, this is not possible in Russian; see (117).

(116) *Ale chc-ę też, [**by** wiedzia-ł-a, że jej ufa-m].*
but want-PRS.1.SG also **COMP** know-PST-F.SG COMP she.DAT trust-PRS.1.SG
'But I also want her to know that I trust her.'

(117) **Ja choč-u, **[by** ona zna-l-a, čto ja ej doverja-ju].*
I.NOM want-PRS.1SG **COND** she.NOM know-PST-SG.F that I.NOM she.DAT trust-PRS.1SG

Because of the synonymy of the forms *że* and *iże*, as well as of the forms *żeby, iżby, ażeby, aby* and *by*, it will suffice to concentrate in the following discussion on the semantic difference between *że* and *żeby*. From a morphological point of view, the fused complementizer *żeby* always requires the past tense form of the complement verb, while *że* can occur with all tenses. These formal restrictions have a semantic background. The fused forms describe situations that are counterfactual, hypothetical or potential. Using these forms, the speaker can express wishes, demands, requests, desires, fears etc. Therefore, the fused COMP-forms co-occur with desiderative predicates, predicates of fearing and manipulative predicates, but not, for example, with predicates of knowledge, immediate perception or propositional attitude. In contrast, the simple form *że* occurs in statements. The following sentences exemplify the semantic difference between these forms.

(118) *Wie-m, [**że** przyjecha-ł].*
know-PRS.1SG **COMP** arrive-PST
'I know that he has arrived.'

(119) *Wie-m, [**że** **by** przyjecha-ł].*
know-PRS.1SG **COMP** **COND** arrive-PST
'I know that he would arrive.'

(120) **Wie-m, [**że-by** przyjecha-ł].*
know-PRS.1SG **COMP-COND** arrive-PST
'I know that he would arrive.'

These examples show that it is not possible to combine a predicate of knowledge with the complementizer *żeby*, because *żeby* marks something that cannot be known by a person. Knowledge is not hypothetical or potential. Therefore, the complementizer *że* must be used. What about the example in (119)? In this sentence, the speaker states that he knows that 'he would come' (*by przyjechał*), e.g.

if he had time. If the fused complementizer *żeby* co-occurs with verbs of saying it triggers a semantic shift and the verb comes to express a demand, as illustrated by (121), which contains two complement clauses.

(121) *Pani prokurator powiedzia-ł-a mi,* [***że-by-m***
Mrs prosecuting.attorney tell-PST-SG.F me **COMP-COND-1SG**
poczeka-ł *i* ***że*** ***dostan-ę*** *za chwilę kopi-ę*].
wait-PST and **COMP** **receive-PRS.1.SG** in a while copy-SG.ACC
'The prosecuting attorney told me that I should wait and that I would get a copy in a while.'

Here, *żebym poczekał* refers to a request and *że dostanę* 'that I would get' to a statement. We can assume that this semantic shift is triggered by the contrast 'indicative vs. conditional mood'. Some predicates allow for both complementizers: with and without the conditional marker; e.g. *bać się* 'be afraid':

(122) *Matka się boi,* [***że-by*** *córc-e* *nie*
mother REFL be.afraid.3SG **COMP-COND** daughter-DAT.SG not
sta-ł-o *się* *coś* *zł-ego*].
happen-PST-N.SG REFL something bad-GEN.SG
'Mother is afraid that something bad might have happened to her daughter.'

(123) *Piotr się boi,* [***że*** *Janusz nie* *podoł-a*
Piotr REFL be.afraid.3SG **COMP** Janusz not cope-PRS.3SG
swo-im *obowiązk-om*].
his-DAT.PL duties-DAT.PL
'Piotr is afraid that Janusz would not cope with his duties.'

4.3 The immediate perception construction

As many other languages, Polish has a complementizer dedicated to complement clauses expressing objects of immediate perception. Karolak (1984: 186) lists 18 verbs denoting visual or acoustic perception allowing for the alternation between *że* and *jak* (*obserwować* 'observe', *podsłuchiwać* 'eavesdrop' and others). As in the case of Russian *kak* (cf. example [30] above), Polish makes use of interrogative pronoun with the original meaning 'how'. Just like *że* it obligatorily occupies the first slot in the complement clause.

(124) *Widzia-ł-em,* [***że*** *przed* *chwilą* *dłużej* *rozmawia-ł* *z* *żon-ą*].
see-PST-1SG **COMP** before moment longer talk-PST with wife-INS.SG
'I saw that a moment ago he had talked to his wife.'

(125) *Wiele razy widzia-ł-e-m,* [***jak*** *podpisywa-ł akt-y.*
several times see-PST-M-1SG **how** sign-PST agreement-ACC.PL
sprzedaż-y nieruchomość-i].
sale-GEN.SG property-GEN.SG
'Several times I saw him signing purchase agreements for properties.'

Whereas *jak* draws attention to the activity or state, the combination of immediate perception predicates with *że* concentrates on the fact.

4.4 Interrogative pronouns and indirect question markers as complementizers

Like in other languages, complement clauses can be introduced in Polish by interrogative elements: the marker of polar questions *czy*, see (126), or interrogative pronouns occurring in WH questions; see (127) and (128). In contrast to its Russian equivalent *li*, which is an enclitic, the Polish marker of polar questions occupies the first slot in the clause.

(126) *Nie wiedzia-ł-e-m nawet,* [***czy*** *umi-e mówi-ć*].
not know-PST-M-1SG ever **Q** be.able-PRS.3SG speak-INF
'I even didn't know whether he was able to speak.'

(127) *Wie-m dobrze,* [***co*** *podoba się publiczność-i*].
know-PRS.1SG well **what** like.prs.3sg REFL public-DAT.SG
'I know quite well what the public likes.'

(128) *A czy nie wie ksiądz,* [***kto*** *móg-ł-by*
what Q not know-PRS.3SG reverent **who** could-PST-COND
nam pomó-c]*?*
us help-INF
'Reverent, don't you know who could help us?'

4.5 Temporal subordinators as complementizers

Just like in Russian (see Section 3.4 above), Polish temporal subordinators show complementizer-like uses; after verbs of liking as e.g. *lubić* 'love' or *podobać się* 'appeal' one can choose between the default complementizer *że* and the temporal subordinator *kiedy* 'when'. The semantic relationship between these complementizers is that both express objective and epistemic modality, but *że* describes actually occurring events, situations or processes, whereas *kiedy* expresses non-actual ones that can be certain, probable or possible.

(129) *Podoba* *mi* *się,* [***że*** *w* *Ameryc-e* *najbogats-i*
like-PRS.3SG me REFL **COMP** in America-LOC.SG richest-PL
konsumenc-i *kupuj-ą* *te* *same* *rzecz-y,* *co* *biedn-i*].
consumer-PL buy-PRS.3.PL these same thing-ACC.PL what poor-PL
'What's great about America is that the richest consumers buy the same things as the poorest.'

(130) *Strona* *rosyjska* *lub-i,* [***kiedy*** *w* *Polsc-e* *jest* *awantura*].
side Russian like-PRS.3SG **when** in Poland-LOC is row.
'The Russians like it when there is a row in Poland.'

(131) *Nie* *podoba* *mi* *się,* [***kiedy*** *dziennikarz* *występ-uje*
not like.PRS.3SG me REFL **when** journalist adopt-PRS.3SG
w *rol-i* *policjant-a*].
in role-LOC.SG policeman-GEN.SG
'I don't like it when a journalist adopts the role of a policeman.'

5 The complementizer system of Bulgarian

The Bulgarian part of the article, as well as the Polish one, will be restricted to an outline of the main properties of complementizers. The Bulgarian system shares one main feature with both Polish and Russian: most complementizers are derived from interrogative pronouns. We will briefly discuss the markers *tova*, *če*; *da*; *kak*; *ako*; *kogato*; *li*; *dali* and interrogative proforms.

5.1 Correlative constructions

In Bulgarian, as well as in Russian, most complementizers have variants with *tova* 'that' which is a correlative functionally similar to the Russian *to* 'that'.[6] However, the statistical distribution of the variants with and without the correlative element is not equal in the two languages. In Russian, the correlative-less variant *nadejus' čto* 'I hope that' is found about 43 million times in Google compared to 5,5 million examples of the variant with the correlative *nadejus' na to čto*. In Bulgarian, 287 thousand examples of the correlative-less varian *nadjavam se če* 'I hope that' correspond to 8 examples of the variant with the correlative *nadjavam se na tova če*. Only 27 example are found for the third person variant with

6 Although there are some exceptions, such as *ako* 'if' and *da* (irrealis complementizer), which are incompatible with the correlative.

the correlative *se nadjava na tova če* '(s/he) hopes that', compared to 1,52 million respective examples without the correlative *tova*. In Russian, in the third person, the respective variant with the correlative *to* even prevails over the correlative-less variant (3,45 million hits with *to* against 950 thousand hits without *to*).

We can state that in general, the variants with the correlative are less widespread in Bulgarian than in Russian. This is obvious not only from the token frequencies, but also from the fact that the range of uses of the Bulgarian correlative *tova* is much narrower than of its Russian counterpart. For example, the correlative is not used in Bulgarian with interrogative pronouns; examples like (132) are not found, though the combination of *tova* and *koj* without the relativizing marker *to* is grammatically possible.

(132) ???*Zna-m* *tova* *koj[to]* *dojde*.
know-PRS.1SG this.N.SG who(REL) come.AOR.3SG
'I know who came.'

For Russian, such examples can be stylistically worse than their corrletaive-less variants, but they nevertheless occur in texts. In Bulgarian, correlatives regularly occur in contrastive contexts only, but even in these contexts the proportion of correlatives is much lower than in Russian.

5.2 The complementizer *če*

The unmarked complementizer for describing real situations in the embedded clause is *če* 'that.' Note that this marker, contrary to the Russian *čto* 'that', does not have interrogative function. In Bulgarian, *kakvo* is used as the interrogative pronoun 'what'. However, historically this marker goes back to the Slavic interrogative pronoun 'what' which is the precursor of the Russian *čto*.

(133) *Ne* *razbira-š* *li* [***če*** *se* *čuvstva-m*
not understand-PRS.2SG whether **that** REFL feel-PRS.1SG
po *sъšti-ja* *način*].
in same-DEF.M.OBL way
'Don't you understand that I feel in the same way?'

(134) *Interesno* *e* [***če*** *tuk* *e* *zapazen* *edinstveni-jat*
interesting be.PRS.3SG **that** here be.PRS.3SG stored sole-DEF.M.NOM
original].
original
'It's interesting that the only original is stored here.'

As in Russian, this marker does not only mark direct objects and subjects, but also peripheral types of arguments. In (135), for instance, it marks an indirect object with the semantic role of Reason/Stimulus.

(135) *Radva -m* *se* [***če*** *se* *vidja-chme*].
be.glad-PRS.1SG REFL **that** REFL see-AOR.1PL
'I am glad that we have met.'

Contrary to Russian and Polish, this complementizer does not form a complex irrealis complementizer with the conditional morpheme (cf. the Russian *čtoby* and the Polish *że-by*). In unreal contexts, the special complementizer *da* is used.

5.3 The complementizer *da*

The complementizer *da* is used when the situation in the embedded clause is not real. For instance, it is widely used with verbs of wish.

(136) *Iska-m* [***da*** *pomogn-a* *na* *novo-to* *rъkovodstvo*].
want-PRS.1SG **COMP** help-PRS.1SG to new-DEF.N management
'I want to help the new management.'

Note that Bulgarian lacks the infinitive. This is why in many (modal) contexts associated in other Slavic languages with infinitives, Bulgarian uses the *da*-construction.

The distribution of *da* is wider than of the Russian *čtoby*. For example, *da* is compatible with some verbs of supposition even without negation; see (137). The Russian *čtoby* is not available in such contexts; see (138).

(137) *V* *tjax* *zlato-to* *se* *predpolaga* [*da* *e*
in they.OBL gold-DEF.N REFL suppose.PRS.3SG COMP be.PRS.3SG
59 *tona*].
59 tons
'In them, it is supposed, there are 59 tons of gold.'

(138) **Predpolagaj-et-sja* [*čtoby* *v* *nix* *by-l-o* *59* *tonn*
suppose-PRS.3SG-REFL that:COND in they.LOC be-PST-SG.N 59 ton.GEN.PL
zolot-a].
gold-SG.GEN
Intended: 'It is supposed that there are 59 tons of gold in them.'

As in Russian, some epistemic verbs hosting embedded clauses can be used with *da* under negation; in this case, the event in the embedded clause loses its realis status:

(139) *Ne si spomnj-am [da sъm vižda-l*
not REFL.DAT remember-PRS.1SG COMP be.PRS.1SG see-PST
togava razbit-i krepost-i].
at.that.time destroyed-PL castle-PL
'I don't remember to have seen destroyed castles at that time.'

This is impossible or marginal in affirmative constructions:

(140) **Spomnj-am si [da sъm vižda-l togava*
remember-PRS.1SG REFL.DAT COMP be.PRS.1SG see-PST at.that.time
razbit-i krepost-i].
destroyed-PL castle-PL
Intended: 'I remember that I have seen destroyed castles at that time.'

Note that the status of *da* has been a subject of controversy and different analyses. According to Nicolova (2008), among others, *da*-forms constitute a special mood with a range of irrealis uses which is usually set apart from indicative mood forms. *Da*-forms can be used in the main clause, just as the subjunctive mood forms in Russian. The decision to regard *da* as a mood marker, and not as a complementizer is problematic from systemic point of view: in general, sentential arguments without a complementizer are much rarer in Bulgarian than in Russian. For instance, (141) sounds very strange.

(141) ???*Misli-x, [ti si zamina-l*].
think-AOR.1SG you REFL.DAT leave-PST
'I thought that you had left.'

Existence of variants where *da* does not occupy the initial position in the complement clause point to its non-complementizer status:

(142) *Ne iska-m [tova da svъrš-i nikoga*].
NEG want-PRS.1SG this COMP end-PRS.3SG never
'I don't want this to end.'

Note that in examples like this, the negative-polarity item *nikoga* is licensed by the negation in the main clause. This seems to be a piece of evidence proving that *da* does not introduce a finite clause. Compare the Russian example with *čtoby* in (143) which is impossible.

(143) **Ja ne xoč-u [čtoby et-o nikogda slučі-l-o-s'*].
I.NOM NEG want-PRS.1SG to this-NOM.SG.N never happen-PST-SG.N-REFL
Intended: 'I don't want this to ever happen.'

5.4 The immediate perception construction

Just like the two other languages, see example (30) for Russian and (125) for Polish, Bulgarian employs the interrogative proadverb 'how' (*kak* in Bulgarian) in contexts of immediate perception. Athough the vailable literature does not provide exhaustive list of the complement taking verbs compatible with this complementizer, it seems that it is more or less restricted to verbs of perception.

(144) *Vižda-l* *sъm* [***kak*** *korab-i* *se* *borj-at* *sъs* *strašn-i*
see-PST be **how** ship-PL REFL struggle-PRS.3PL with terrible-PL
bur-i].
storm-PL
'I have seen ships struggling with terrible storms.'

In the same or similar contexts, the unmarked complementizer *če* is also allowed:

(145) *Vidja-x* [***če*** *korab-i* *se* *borj-at* *sъs* *strašn-i* *bur-i*].
see-AOR.1SG **that** ship-PL REFL struggle-PRS.3PL with terrible-PL storm-PL
'I saw that ships were struggling terrible storms.'

The semantic distinction between the two variants is the same as in the two other languages: the variant with *kak* emphasizes the process itself (in [144] it is the process of struggling the storms), while the construction with *če* in (145) focuses on the fact of struggling (see Arutjunova 1988 on the distinction between 'processes/events' and 'facts' in linguistics), where the speaker might have observed the situation, but it can be equally the case that he has only perceived some relevant subparts of it.

It is important that the complementizer *kak* in examples like (144) occurs in its 'interrogative' form, without the element *to*, whereas the variant *kakto* is reserved for relative and adverbial contexts. This means that the perception use of *kak* goes back to indirect questions, as in (146), and further to direct questions, as in (147).

(146) *Znam* [***kak*** *raboti* *tazi* *sistema*].
know-PRS.1SG **how** work.PRS.3SG this system
'I know how this system works.'

(147) ***Kak*** *ste?*
how be.PRS.2PL
'How do you do?

5.5 Interrogative pronouns as complementizers

As in the other languages, interrogative pronouns can be used in Bulgarian in sentential arguments, marking complementation. Consider the following examples with *koj* 'who' and *kak* 'how'.

(148) *Zna-m* [***koj*** *e* *neformalni-jat* *lider*].
know-PRS.1SG **who** be.PRS.3SG informal-DEF.M.NOM leader
'I know who is the informal leader.'

(149) *Ne* *razbra* [***kak*** *stigna* *do* *nego*].
not understand.AOR.3SG **how** reach.AOR.3SG to he.OBL
'He didn't understand how she got to him (found him).'

Importantly, in these contexts, the pronouns are used in their interrogative variant, and not in their variant with *-to* (see Section 5.7 below about the use of the adverbial marker *kogato* 'when' which does have *to* in complemention context). The variants with *to* (*kogato*) are generally reserved for relativization (as in *vremeto kogato* 'the time when') and adverbialization (*dojde kogato* '(s/he) came when'). Perhaps this makes indirect questions in argument position closer to questions proper than relative and circumstantial clauses.

Note also that the indirect question use of complementizers should be distinguished from their use in the immediate perception construction. The same marker *kak* 'how' can be used in both functions. However, in the indirect question construction its semantic function is the same as in questions proper: it denotes the 'mode', the 'way' in which the situation takes place (in example [149], for instance, it is the way in which she found him). In contrast, in the immediate perception construction, *kak* serves as a complementizer without the 'mode' component. The difference becomes obvious when we try to introduce modifiers like *točno* 'precisely' into sentences with *kak*. The insertion of *točno* is possible in the indirect question structures and impossible in immediate perception ones.

5.6 The indirect question markers *li* and *dali*

The marker *li*, a cognate of Russian *li*, is used as a complementizer also in Bulgarian. However, its distribution is not the same as in Russian. In literary Bulgarian, all direct yes/no-questions contain *li* (which is a rather old-fashioned strategy for marking such in contemporary Russian). The main device for indirect ques-

tion marking in Bulgarian, functionally comparable to the Russian *li*, is *dali*[7] 'if, whether', though indirect questions with *li* are not infrequent in the Internet either:

(150) *Popita-x* *go* [***dali*** *te* *e* *vižda-l*].
ask-AOR.1SG he.ACC **COMP** you.ACC be.PRS.3SG see.PRF.3SG
'I asked him if he has seen you.'

(151) *Popita-x* *go* [*ima* ***li*** *njakakv-i* *drug-i* *defekt-i*].
ask-AOR.1SG he.ACC have.3SG **COMP** some-PL other-PL defect-PL
'I asked him if there were any other defects.'

The boundary between indirect and direct questions is vague, but in (150) and (151) we are dealing with indirect questions rather than with direct ones. This is clear from the absence of an interrogation mark. Moreover, *dali* and *li* occur in indirect questions with verbs which do not denote a question, such as *interesuvam se* 'be interested'. The range of verbs allowing *li* seems to be narrower than the range of verbs allowing *dali*. For instance, the verb *bespokoja se* 'worry' can be used with *dali* but not with *li*. With deverbal nouns the usage of *li* is even more restricted. For instance, with the noun *vъpros* 'question' no examples with *li* are found in the Internet, and only two such examples are found in the Bulgarian National Corpus:

(152) *Na* *vъpros* [*ima* ***li*** *dokazatelstv-a* *za* *koristn-i*
on question have.PRS.3SG **whether** proof-PL for mercenary-PL
cel-i *ot* *stran-a* *na* *Aleksandrov*]
aim-PL from side-SG on Alexandrov
'On the question if there are proofs of Alexandrov's mercenary aims ...'
('the minister claimed that on his opinion, there are enough of them.')

As in Russian, the marker *ili* also serves as an indirect question marker. However, it never occurs without the marker *li* or *dali*.

(153) *Ne* *znam* [***dali*** *e* *sъn* ***ili*** *dejstvitel'nost'*].
not know **whether** be.PRS.3SG dream **or** reality
'I don't know whether it is a dream or the reality.'

(154) *Ne* *znam,* [*az* ***li*** *bъrka-m* ***ili*** *taka* *si* *rabot-i*].
not know I.NOM **whether** be.mistaken-PRS.3SG **or** so REFL.DAT work-PRS.3SG
'I don't know whether I do something wrong or it works (itself) like this.'

7 *Dali* is diachronically analyzed as a combination of the general irrealis marker *da* and the polar question marker *li*.

This means that *ili* must be regarded here as a usual conjunction. There are no uses in Bulgaria comparable to Russian examples like (40) where the occurrence of *ili* makes the use of the indirect question marker optional.

5.7 Temporal and conditional subordinators as complementizers

Bulgarian has the same use of the conditional subordinator *ako* 'if' and the temporal one *kogato* 'when' as in Russian:[8]

(155) *Ne obič-am* [***kogato*** *mi govorj-at kato na bebe*].
not love-PRS.1SG **when** I.DAT talk-PRS.3PL like to baby
'I don't like it when people talk to me as to a baby.'

(156) *Popita-j go šte mu charesva li* [***ako*** *go otbjagva-š*].
ask-IMV he.ACC FUT he.DAT like.PRS.3SG Q **if** he.ACC avoid-PRS.2SG
'Ask him if he will like it if you avoid him.'

It seems that the same restriction on the word order is applicable, as in Russian: when adverbial subordinators are used as complementizers, the matrix clause always precedes the embedded one. This is why (157) is ungrammatical:

(157) [**koga.to mi govorj-at kato na bebe*], *ne obič-am*
when I.DAT talk-PRS.3PL like to baby not love-PRS.1SG
Intended: 'I don't like it when people talk to me as to a baby.'

The statistical distribution of these constructions in Russian and Bulgarian is different. In Bulgarian, they seem to be more restricted; some cognates of verbs that allow "adverbial" marking of clausal complements in Russian disprefer such marking in Bulgarian: this is true for *bezpokoja* 'worry.' No examples with this verb and a sentential argument with *kogato* 'when' used as a complementizer are found in the Internet and only one example is found in the Bulgarian National Corpus:

(158) *Izgležda vi bezpoko-i* [***kogato*** *čovek njama*
appear.PRS.3SG you.ACC worry-PRS.3SG **when** person have.NEG.PRS.3SG
griži].
concerns
'It seems that you worry if a person doesn't have any concerns.'

8 In Bulgarian, interrogative pronouns like *koga* 'when' get the particle *to* when used as adverbial or relative subordinators.

5.8 Complementizer omission

In Bulgarian complement clauses without a complementizer are rarer than in Russian. However, this issue calls for more detailed investigation.

6 Conclusion

In our contribution, we have tried to give an account of the COMP-systems of Russian, Polish and Bulgarian. The analysis was based on the assumption that complementizers are syntactic words traditionally labelled as 'conjunctions'. We have shown that complementizers can have semantic functions, mainly in the field of epistemicity and evidentiality, or in the case of interrogative elements, in the field of indefiniteness. We would like to leave for further discussion the question how we can distinguish between complementizers in a proper sense and sentence adverbs and adverbial subordinators. There is evidence that COMP is strongly associated with the left periphery of the clause and might be grasped in terms of a specific syntactic slot preceding the core predication. This fact yields an interesting situation where there are several units with intermediate properties between complementizers and particles. Some of them pass tests for being complementizers, while others do not, though their association with the initial position in the embedded clause shows that they are on their way to become complementizers.

An interesting feature of Russian is its rich system of composed markers (those with the correlative *to* or with the subjunctive marker *by*) which in fact can be analyzed either as complex complementizers (this is more plausible for *by*-variants) or as a combinations of markers (this seems to be plausible for *to*-variants).

Russian has no specialized complementizers, such as the Polish *że* having only a complementizer use. All Russian markers discussed in the article have other uses as well. Sometimes, as with the polysemous adverbial clause markers *kogda* 'when', *esli* 'if' and the purpose marker *čtoby* '(in order) to', the word order can serve as a distinctive feature: in their use as complementizers, these units (and, correspondingly, the embedded clause itself) have a fixed position after the main clause, which is not the case with adverbial uses, which can be also preposed in relation to the main clause. This fact provides support (at least for *kogda* and *esli*) for the claim that if a marker is used as a complementizer and as another type of subordinator, the complementizer use is often secondary, a result of gram-

maticalization (see Pekelis 2014, who claims that complementizers are semantically more abstract and vague than adverbial subordinators).

To sum up, we have shown that the complementizer systems of Russian, Polish and Bulgarian express the following functional contrasts:

- 'object of immediate perception' vs. 'knowledge acquired';
- degrees of epistemic distancing;
- epistemological contrast: indefiniteness contrasts; i.e. contrasts related to speaker's knowledge ('speaker knows *p*' vs. 'speaker does not know *p*');
- 'assertion, informative utterance' vs. 'utterance of a demand';
- aspectual contrasts with verbs of liking; the alternation between the standard unmarked complementizer and temporal or conditional subordinators can mark the distinction between single and non-episodic events;
- obviation effects in same subject constructions involving different degrees of speaker control over the event;
- in Russian and Polish: mental act of the speaker (first person marking) vs. mental act of a non-participant in the communicative event (third person marking);
- in Bulgarian: 'realis' vs. 'irrealis' distinction;
- in Polish: stylistic contrasts ('neutral' vs. 'bookish').

We assume that these contrasts are not exclusively marked by the complementizer itself, but are rendered by the construction as a whole, involving features like the semantics of the predicate, the mood choice and argument sharing. It has been shown that mood, especially the dichotomy between infinitive, subjunctive and indicative, plays a major role in the encoding of complement clauses. The internal make-up of complement clauses cannot be accounted for if we do not take into consideration the syntactic mechanisms of control and obviation.

Acknowledgment

This article was prepared within the framework of the Academic Fund Program at the National Research University Higher School of Economics (HSE) in 2015–2016 (grant no. 15-01-0150) and supported within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

References

Arutjunova, Nina D. 1988. *Tipy jazykovyx značenij. Ocenka. Sobytie. Fakt* [Types of linguistic meanings. Evaluation. Events. Facts.]. Moscow: Nauka.

Avrutin, Sergey & Maria Babyonyshev. 1997. Obviation in subjunctive clauses and AGR: Evidence from Russian. *Natural Language and Linguistic Theory* 15. 229–262.

Baker, Mark C. & Nadya Vinokurova. 2009. On agent nominalizations and why they are not like event nominalizations. *Language* 85(3). 517–556.

Barnetová, Vilma, Helena Běličová-Křížková, Oldřich Leška, Zdena Skoumalová & Vlasta Straková (eds.). 1979. *Russkaja Grammatika 2* [Russian Grammar 2]. Praha: Academia.

Belošapkova, Vera A. 1977. *Sovremennyj russkij jazyk: Sintaksis* [The modern Russian language: Syntax]. Moscow: Vysšaja Škola.

Boye, Kasper. 2010a. Reference and clausal perception-verb complements. *Linguistics* 48(2). 391–430.

Boye, Kasper. 2010b. Evidence for what? Evidentiality and scope. In Björn Wiemer & Katerina Stathi (eds.), *Database on evidentiality markers in European languages*. [Special issue]. STUF – Language Typology and Universals 63(4), 290–307. Berlin: Akademie-Verlag.

Boye, Kasper. 2012. *Epistemic meaning: A crosslinguistic and functional-cognitive study*. Berlin & Boston: De Gruyter Mouton.

Brecht, Richard R. 1977. *Čtoby* or *Čto* and *By*? *Folia Slavica* 1(1). 33–41.

Bresnan, Joan. 1970. On complementizers: Toward a syntactic theory of complement types. *Foundations of Language* 6. 297–321.

Comrie, Bernard. 1971. *Aspects of sentence complementation in Russian*. Cambridge: King's College dissertation.

Decyk-Zięba, Wanda & Stanisław Dubisz. 2008. *Glosariusz staropolski: Dydaktyczny słownik etymologiczny* [Glossary of Old-Polish: Didactic etymological dictionary]. Warszawa: WUW.

Dixon, Robert M. W. 2006. Complement clauses and complementation strategies in typological perspective. In Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation: A cross-linguistic typology*, 1–48. New York: Oxford University Press.

Dobrushina, Nina. 2012. Subjunctive complement clauses. *Russian Linguistics* 36(2). 121–156. http://link.springer.com/article/10.1007/s11185-012-9088-0/fulltext.html (accessed June 2014).

Farkas, Donka. 1988. On obligatory control. *Linguistics and Philosophy* 11(1). 27–58.

Franks, Steven & Tracy Holloway King. 2000. *A handbook of Slavic clitics*. Oxford: Oxford University Press.

Hansen, Björn. 2010. A constructional approach to the gradual rise of epistemic sentence adverbs in Russian. In Björn Hansen & Jasmina Grković-Major (eds.), *Diachronic Slavonic syntax: Gradual changes in focus* (Wiener Slawistischer Almanach – Sonderband 74), 75–87. München, Berlin & Wien: Otto Sagner.

Karolak, Stanisław. 1984. Składnia wyrażeń predykatywnych [Syntax of predicative expressions]. In Zuzanna Topolińska (ed.), *Gramatyka współczesnego języka polskiego. Składnia* [Grammar of contemporary Polish], 11–211. Warszawa: PWN.

Khomitsevich, Olga. 2007. *Dependencies across phases*. Utrecht: University of Utrecht PhD. Dissertation.

Knjazev, Michail Ju. 2009. *Predikatnyje aktanty s mestoimeniem to v russkom jazyke: grammatičeskie osobennosti i osobennosti upotreblenija* [Argument clauses with the

pronoun *to* in Russian: grammatical features and use]. Saint-Petersburg: Saint-Petersburg University Diploma work.

Kobozeva, Irina M. 2011. Uslovija upotreblenija to pered pridatočnym izjasnitel'nym [Conditions of use of *to* before a complement clause]. Materials from the conference Dialog-2001. Moscow: RGGU.

Kobozeva, Irina M. 2012. Uslovija upotreblenija *to* pered pridatočnym iz'jasnitel'nym s sojuzom *čto* [Conditions of use of *to* before a complement clause with the complementizer *čto*]. In Olga Inkova (ed.), *Du mot au texte. Etudes slavo-romanes*, 129–148. Bern: Peter Lang.

Landau, Idan. 2000. *Elements of control. Structure and meaning in infinitival constructions*. (Studies in Natural Language and Linguistic Theory 51). Dordrecht, Boston & London: Kluwer.

Letuchiy, Alexander B. 2014. Russian *esli*-constructions: arguments or adjuncts? Factive or non-factive? Talk at 36. Jahrestagung der Deutschen Gesellschaf für Schprachwissenschaft, Marburg, 5–7 of March 2014.

Letučij, Aleksandr. 2008. Sravnitel'nye konstrukcii, irrealis i ėvidencial'nost'[Comparative constructions, irrealis and evidentiality]. In Björn Wiemer & Vladimir A. Plungjan (eds.), *Lexikalische Evidenzialitätsmarker in slavischen Sprachen*, 215–238. München: Sagner.

Letučij, Aleksandr. 2012. O nekotoryx svojstvax russkich sentencial'nych aktantov [On some properties of Russian argument clauses]. *Voprosy jazykoznanija* 5. 57–87.

Markowski, Andrzej. 2000. *Nowy słownik poprawnej polszczyzny* [A new dictionary of correct Polish]. Warszawa: PWN.

Meyer, Roland. 2004. *Syntax der Ergänzungsfrage. Empirische Untersuchungen am Russischen, Polnischen und Tschechischen*. München: Sagner.

Nicolova, Ruselina. 2008. *Bъlgarska gramatika – morfologija* [Bulgarian grammar: morphology]. Sofija: Sv. Kliment Oxridski.

Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description 2*: *Complex constructions*, 2nd edn., 52–150. Cambridge: Cambridge University Press.

Nuyts, Jan. 2001. *Epistemic modality, language and conceptualization*. Amsterdam: Benjamins.

Padučeva, Elena V. 2010. Faktivnost' i kosvennyj vopros [Factivity and indirect question]. In Atle Grønn & Olga Klonova (eds.), *Russkij jazyk v kontrastivnom aspekte* [The Russian language in contrastive perspective] (Oslo Studies in Language 2[3]), 479–494. Oslo.

Pekelis, Olga. 2014. Infinitiv vs. pridatočnoe s sojuzom čtoby: k voprosu o vybore sposoba oformlenija sentencial'nogo aktanta v russkom jazyke [Infinitives vs. dependent clauses with the complementizer *čtoby*: On the choice of the strategy of marking argument clauses in Russian]. *Voprosy jazykoznanija* 4. Moscow: Nauka.

Ramat, Paolo & Davide Ricca. 1998. Sentence adverbs in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 187–277. Berlin: Mouton de Gruyter.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar – Handbook of Generative Grammar*, 281–337. New York: Springer.

Serdobol'skaya, Nataliya V. 2011. K tipologii vyraženija generičeskogo sobytija v konstrukcijach s sentencial'nymi aktantami [Towards a typology of expression of generic events in constructions with argument clauses]. Handout at the 8th Conference in Typology and Grammar for Young Researchers. Saint-Petersburg.

Švedova, N. Ju. (ed.). 1980. *Russkaja grammatika (t. II): Sintaksis* [Russian grammar, volume II: Syntax]. Moscow: Nauka.

Szucsich, Luka. 2010. Obviation und temporale Abhängigkeit bei Subjunktiven. *Zeitschrift für Slawistik* 54(4). 398–415.

Testelec, Jakov G. 2001. *Vvedenie v obščij sintaksis* [Introduction to general syntax]. Moscow: Rossijskij Gosudarstvennyj Gumanitarnyj Universitet.

Valgina, Nina S. 2000. *Sintaksis sovremennogo russkogo jazyka* [Syntax of modern Russian]. Moscow: Vysšaja Škola.

Wiemer, Björn. 2008. Pokazateli s citativnoj i inferentnoj funkcijami [Markers with hearsay and inferential functions]. In Björn Wiemer & Vladimir A. Plungjan (eds.), *Lexikalische Evidenzialitätsmarker in slavischen Sprachen*, 335–376. München: Otto Sagner.

Zaliznjak, Anna A. & Irina L. Mikaelyan 1988. Ob odnom sposobe razrešenija neodnoznačnosti v glagolach propozicional'noj ustanovki [On one way of disambiguation of constructions with verbs of propositional attitude]. In *Semantičeskie aspekty formalizacii intellektual'noj dejatel'nosti* [Semantic aspects of formalizing intellectual activity]. Moscow: VINITI RAN.

Zimmermann, Ilse. 2009. Satzmodus. In: Peter Kosta, Tilman Berger, Karl Gutschmidt & Sebastian Kempgen (eds.), *Slavic languages. An international handbook of their history, their structure and their investigation. Vol 1.* (Handbücher zur Sprach- und Kommunikationswissenschaft 32.1), 484–509. Berlin & New York: Walter de Gruyter.

Axel Holvoet

Semantic functions of complementizers in Baltic

1 Introductory remarks

The article starts with some general information on the languages dealt with, Lithuanian and Latvian. Then an overview of the basic Baltic complementizers and their distributional properties is given. It is followed by an overview of more specialized complementizers, occurring with specific types of complement-taking predicates, or competing with the more basic complementizers when some additional meaning shades are to be conveyed. This leads, in a natural way, to a discussion of the semantic functions of complementizers. The discussion concentrates on three types of distinctions, namely between propositional vs. state-of-affairs, realis vs. irrealis, and epistemically neutral vs. epistemically marked complementizers. It is argued that these three distinctions can basically function independently of each other, though they may also coincide by default. Attention is also given to historical developments as well as to polyfunctionality, omission and combination of complementizers.

1.1 Notes on the Baltic languages

Lithuanian and Latvian are the only living representatives of the Baltic subgroup of Indo-European languages, which many hold to be so closely related to Slavonic as to entitle us to posit a Balto-Slavonic subgroup. Both Lithuanian and Latvian are attested in writing from the 16th century onward, but the standard languages did not acquire their definitive shape until the early 1920s, when independent Baltic States were established. Before that time, they were largely restricted to the sphere of the everyday life of the rural communities. Throughout the 16th–18th centuries, the written varieties were used mainly in religious literature; Old Lithuanian and, still more, Old Latvian writings were heavily dependent on foreign models (Polish and German respectively), especially in syntax. In Latvia, religious differences between Lutheran Courland and Swedish Livonia and Catholic Polish Livonia led to the rise of two written varieties: Standard Latvian and Latgalian. Latgalian is still used as a regional written standard, mainly in the sphere of activities of the Roman Catholic Church, and is sometimes described as a separate language. The national revival movements led to the rise of original

Lithuanian and Latvian literatures in the 19th century. Lithuanian is now spoken by about 2,900,000 people in Lithuania and some 300,000 abroad, Latvian by some 1,380,000 people in Latvia and 120,000 abroad. Lithuanian and Latvian are fusional languages with rich case systems and elaborate systems of verbal morphology abounding in participles and converbs. Lithuanian has always had close areal links to neighbouring Slavonic (Old Russian, Belorussian, Polish, Russian) and bears traces of this in vocabulary and syntax. Latvian developed largely on a Fennic (Livonian) substratum (Livonian survived into the 20th century in a few coastal villages west of the Bay of Riga) and the stamp of Fennic influence is clearly felt in Latvian phonetics and syntax. Latvian has moreover always been exposed to a strong German influence, with marked influences on vocabulary and syntax. Under Tsarist rule and again in Soviet times the influence of Russian was strong and came to outweigh that of Polish on Lithuanian and that of German on Latvian. Old Prussian, the language of the autochthonous population of East Prussia, is attested in a few 16th century texts of rather poor quality. It became gradually extinct in the 17th century. The present study is confined to modern standard Lithuanian and Latvian.

1.2 Complementation strategies in Baltic

Both Lithuanian and Latvian have several competing complementation strategies, and finite clauses with complementizers are only one of them. Infinitival complements naturally occur with phasal, manipulative and achievement (implicative) predicates:

Lithuanian[1]

(1) *Vaik-ai* *pradėjo* [*žaisti* *slėpyni-ų*].
child-NOM.PL start.PST.3 play.INF hide.and.seek-GEN.PL
'The children started playing hide-and-seek.'

A characteristic feature of Baltic is the widespread use of converbs and participles in complementation. Converbs are used in constructions with verbs of perception and less often with utterance verbs and verbs of propositional attitude:

1 Part of the examples in this article are constructed, especially for the purpose of providing minimal pairs. Most examples are, however, authentic. For examples taken from 20th century fiction only the writer's name is given, whereas for examples taken from the Internet an URL is provided. With regard to the glossing it should be noted that 3rd person finite verb forms never distinguish number in Baltic, so that these forms are simply glossed '3'.

Latvian

(2) *Es* *dzirdēju* [*Jān-i* *dzied-am*].
1SG.NOM hear.PST.1SG John-ACC sing-CVB
'I heard John singing.'

Declinable participles are used with utterance, knowledge and propositional attitude predicates, but only in the case of coreferential subjects:

Lithuanian

(3) *Marij-a* *sakė-si* [***to*** *ne-žinojusi*].
Mary-NOM say.PST.3-REFL **that** NEG-know.PPA.NOM.SG.F
'Mary said she hadn't known this.'

In the modern standard varieties, however, complements with converbs and participles are becoming less frequent (they are mostly used in formal style), and finite complements with complementizers have imposed themselves (alongside infinitival constructions) as the predominant complementation strategy.

2 Overview of complementizers in Baltic

2.1 Basic complementizers

In this section I will deal with those units (i) for which the function of complementizer in Noonan's ([1985] 2007) sense is basic (in contrast to adverbial subordinators occasionally used in complementation), (ii) that are most general in scope (that is, they are not restricted to one type of complement-taking predicates), and (iii) that can be regarded as unmarked with regard to other units (that is, they can be optionally replaced with other units with specific types of complement-taking predicates and for the purposes of specific types of additional semantic marking).

Contemporary Lithuanian works with two basic complementizers: *kad* is epistemically neutral and potentially factive if the complement-taking predicate suggests a factive interpretation, whereas *ar* marks interrogation or suspension of judgement and is always non-factive (Bresnan's "undetermined meaning" (Bresnan 1979: 27). *Kad* will be used with predicates of knowing and acquisition of knowledge, with predicates of propositional attitude and with utterance predicates. *Ar* is used, for example, with speech act verbs like *klausti* 'ask', but also with verbs associated with acquisition of knowledge such as *patikrinti* 'check', *pasitikslinti* 'ascertain' etc.:

Lithuanian

(4) a. *Jon-as* *sakė,* [***kad*** *būsi* *koncert-e*].
John-NOM said.PST.3 **COMP** be.FUT.2SG concert-LOC.SG
'John said you would be at the concert.'

Lithuanian

(4) b. *Jon-as* *klausė,* [***ar*** *būsi* *koncert-e*].
John-NOM ask.PST.3 **COMP** be.FUT.2SG concert-LOC.SG
'John asked if you would be at the concert.'

But *ar* is not restricted to interrogative complements: it is also used as a non-factive complementizer with verbs of knowing and saying, especially within the scope of negation:

Lithuanian

(5) *Jon-as* *ne-sakė,* [***kad*** *eis* *į* *koncert-ą*].
John-NOM NEG-say.PST.3 **COMP** go.FUT.3 to concert-ACC.SG
'John didn't say he would be going to a/the concert.'

Lithuanian

(6) *Jon-as* *ne-sakė,* [***ar*** *eis* *į* *koncert-ą*].
John-NOM NEG-say.PST.3 **COMP** go.FUT.3 to concert-ACC.SG
'John didn't say whether he would be going to a/the concert.'

Latvian has three instead of two basic complementizers. The complementizer used for interrogation and suspension of judgement is *vai*:

Latvian

(7) *Jān-is* *ne-teica,* [***vai*** *viņ-š* *atnāks* *uz* *izstād-i*].
John-NOM NEG-say.PST.3 **COMP** 3-NOM.SG.M come.FUT.3 to exhibition-ACC
'John did not say whether he would come to see the exhibition.'

The field covered by Lithuanian *kad* is divided between two distinct complementizers: *ka* is basically used in propositional (truth-valued) complements whereas *lai* is used in state-of-affairs complements (this distinction corresponds to that between "third-order" entity and "second-order entity" in Lyons (1977: 443–445), "truth modality" and "action modality" in Ransom (1986: 5–6), "fact type" and "potential type" in Dixon (2006: 24–25) etc.; for a recent overview of terminology, with discussion, cf. Boye (2010). While *ka* is characteristic of complements

with verbs of knowledge and propositional attitude as well as with utterance predicates, *lai* is characteristic of desiderative predicates like 'want', 'wish', and of utterance predicates reflecting acts of volition, such as 'ask', 'request', 'demand' etc.

Latvian

(8) a. *Es* *zinu,* [***ka*** *tu* *atbrauksi* *paciemoties*].
1SG.NOM know.PRS.1SG **COMP** 2SG.NOM come.FUT.2SG visit.INF
'I know you will come on a visit.'

Latvian

(8) b. *Es* *gribu,* [***lai*** *tu* *atbrauc* *paciemoties*].
1SG.NOM want.PRS.1SG **COMP** 2SG.NOM come.PRS.2SG visit.INF
'I want you to come on a visit.'

It should be noted that 'hope' belongs to the propositional type:

Latvian

(9) *Es* *ceru,* [***ka*** *man* *tā* *ne-gadīsies*].
1SG.NOM hope.PRS.1SG **COMP** 1SG.DAT so NEG-happen.FUT.3
'I hope this will not happen to me.'

This treatment is evidently connected, in turn, with the strong epistemic component in the meaning of 'hope', which combines a desiderative element with a high degree of confidence in the realization of the desirable state-of-affairs. However, Noonan's classification (Noonan 2007: 132) of desiderative predicates into a 'hope' and a 'wish' type, the latter being assumed to be inherently counterfactive, does not seem to fit the facts. Desiderative predicates are inherently non-factive, that is, they entail neither realization nor non-realization of the desirable state-of-affairs. Of course, there may be different *expectations of actuation* (a term here borrowed from Mithun 1995), and these may receive distinctive marking. In the case of 'hope' the likelihood of actuation tends to be reflected in the selection of a propositional complement type. At the other end of the spectre, low expectation of actuation may be emphasized by means of irrealis marking, as we will see further on.

The choice between the complemetizers *ka* and *lai* is often dictated by the complement-taking verb, but with speech act verbs like *teikt* and *sacīt* 'say' and other verbs conveying transfer of information, such as *rakstīt* 'write', the choice of complementizer will disambiguate between a propositional and a state-of-affairs (in fact, volitional) complement:

Latvian

(10) a. *Jān-is* *rakstīja,* [***ka*** *mums* *atsūtīs* *naud-u*]. L
John-NOM write.PST.3 **COMP** 1PL.DAT send.FUT.3 money-ACC
'John wrote that he would send us some money.'

Latvian

(10) b. *Jān-is* *rakstīja,* [***lai*** *atsūtām* *viņ-am* *naud-u*].
John-NOM write.PST.3 **COMP** send.PRS.1PL 3-DAT.SG.M money-ACC
'John wrote [asking] we should send him some money.

In Lithuanian the same would be possible, but here the difference would be conveyed by mood only, the irrealis (conditional) being selected to obtain a volitional interpretation (note that there is no difference in mood between [10a] and [10b] in Latvian):

Lithuanian

(11) a. *Jon-as* *sakė,* [***kad*** *pasitiks* *mus* *stot-yje*].
John-NOM say.PST.3 **COMP** meet.FUT.3 1PL.ACC station-LOC.SG
'John said he would meet us at the station.'

Lithuanian

(11) b. *Jon-as,* *sakė,* [***kad*** *pasitiktume* *jį* *stot-yje*].
John-NOM say.PST.3 **COMP** meet.IRR.1PL 3.ACC.SG.M station.LOC.SG
'John told us to meet him at the station.'

In Latvian, the extent to which the choice of the complementizer – *ka* or *lai* – may have a semantic import of its own and determine the interpretation of the clausal complement is variable. In a pair of sentences like (10a) and (10b) the complementizer certainly disambiguates between two types of complements that can occur with speech act verbs, namely complements expressing assertive and directive speech acts. But the complementizer may display an even heavier functional load, especially with verbs of epistemic stance like *domāt* 'think'. With such verbs, which normally occur with *ka*, the choice of *lai* will entail a deontic reading, and the complementizer may be said to represent a deontic modal predicate:

Latvian

(12) a. *Ak tad tu domā, [**ka** es ar labdarīb-u nodarbojos]?*
oh so 2SG.NOM think.PRS.2SG **COMP** 1SG.NOM with philanthropy-ACC be.occupied.PRS.1SG
'Oh, so you think I engage in philanthropy then?'

Latvian

(12) b. *Ak tad tu domā, [**lai** es ar labdarīb-u nodarbojos]?*
oh so 2SG.NOM think.PRS.2SG **COMP** 1SG.NOM with philanthropy-ACC be.occupied.PRS.1SG
'Oh, so you think I should engage in philanthropy then?'
(Pāvils Rozītis, 1889–1937)

Though the deontic (obligative) reading of *lai* will be a default if the complement-taking predicate is of the type normally requiring a propositional complement, this value is not a constant feature. This complementizer also occurs with verbs that have no volitional element, like *gaidīt* 'wait':

Latvian

(13) *Viņ-a pieliecās At-im, it kā gaidī-dam-a, [**lai** viņ-š to noglāsta].*
3-NOM.SG.F lean.PST.3 Atis-DAT as.if wait-CVB-F **COMP** 3-NOM.SG.M 3-ACC.SG stroke.PRS.3
'She leant towards Atis as if she was waiting for him to stroke her.'
(Kārlis Zariņš, 1889–1978)

The complementizer *lai* is therefore not inherently volitional: it merely represents the state-of-affairs complement type. At the borderline between propositional and state-of-affairs complementation we find manipulative predicates such as *atļaut* 'allow'. This verb often has a non-finite (infinitival) complement, but if the clausal complement is finite, it can be introduced both by *ka* and by *lai*:

Latvian

(14) *Nu kur-š normāl-s vīriet-is atļauj, [**ka** regulāri trauk-us plēš un vēl paš-u dauza]?*
now which-NOM.SG.M normal-NOM.SG man-NOM.SG allow.PRS.3 **COMP** regularly dish-ACC.PL break.PRS3 and even self-ACC.SG pound.PRS.3
'Which normal man would allow dishes regularly to be broken and himself to be pounded?'(http://www.atceries.lv/lv/diskusijos.zinutes/49741, accessed Sept. 2012)

Latvian

(15) *Zviedrij-as* *policij-a* *reizi* *gad-ā* *atļauj,* [***lai*** *bļauj*
Sweden-GEN police-NOM once year-LOC.SG allow.PRS.3 **COMP** shout.PRS.3
un *dūr-es* *vicina* *div-as* *aktīvist-u* *grup-as*].
and fist-ACC.PL shake.PRS.3 two-NOM activist-GEN.PL group-NOM.PL
'Once a year the Swedish police allow two groups of activists to shout and shake their fists.'
(http://www.apollo.lv/zinas/kas-maksas-par-grautinu/398597, accessed Sept. 2012)

This indeterminacy is understandable, as with a verb like 'allow' the complement will be factive in some cases, e.g., in the past tense (*allowed him to enter*) but non-factive in other cases, say, in the imperative (*allow me to come in*). This makes it possible for the complement to be interpreted as belonging either to the propositional or to the state-of-affairs type. There is also variation as to whether the manipulation is effected through the performance of directive speech acts or not. The use of *atļaut* as a speech act verb in a wider sense will cause it to display the type of complementation normally chosen with volitional predicates, that is, the state-of-affairs type, whereas non-verbal manipulation will show a bias towards infinitival complementation, but with the propositional complementizer strategy as an alternative in factive contexts. Thus several factors interact in determining the choice of the type of complement.

2.2 Other words introducing clausal complements

Clausal complements can also be introduced by question words like Lith. Latv. *kas* 'who, what', Lith. *kaip*, Latv. *kā* 'how' etc. These are not formally distinguished from interrogative and relative pronouns. As in many other languages (e.g., English), there is thus often no formal difference between an interrogative complement clause and a free relative clause. In the following pair of examples, only tense distinguishes the two (the interrogative complement displaying lack of tense backswitching and therefore containing a present tense marking simultaneity with regard to the matrix clause) whereas the relative clause has a past tense parallel to that of the matrix clause:

Latvian

(16) a. *Prezident-s* *ne-pateica,* [*ko* *domā*].
president-NOM.SG NEG-say.PST.3 what.ACC think.PRS.3
'The president didn't say what he thought.' (complement clause)

Latvian

(16) b. *Prezident-s* *ne-pateica,* [*ko* *domāja*].
president-NOM.SG NEG-say.PST.3 what.ACC think.PST.3
'The president didn't say (that which) he thought' (free relative clause)

While the use of question words in interrogative clausal complements is a widespread feature, Lithuanian has a highly specific development consisting in the question word *kam* 'why, to what purpose' losing its interrogative status and developing into a specialized complementizer with such verbs as *priekaištauti* 'reproach', *kritikuoti* 'criticize' etc.:

Lithuanian

(17) *Sapn-e* *vien-ai* *moteri-ai* *laum-ė* *priekaištavo,*
dream-LOC.SG one-DAT.PL.F woman-DAT.PL fairy-NOM.SG reproach.PST.3
[*kam* *ši* *jos* *vaik-ą* *plaka*].
why DEM.NOM.SG.F 3.GEN.SG.F child-ACC.SG beat.PRS.3
'In a dream a fairy reproached one woman with beating her [the fairy's] child.'
(http://www.aruodai.lt/paieska/objektas.php?Oid=2783, accessed Sept. 2012)

The meaning was originally 'reproached her, [asking] why she beat her child'. This process, which may eventually lead to the rise of a new complementizer, dissociated from the original question word, is not yet completed.

3 Less basic complementizers

3.1 Immediate perception

Immediate perception predicates share the same complementation strategy in Lithuanian and Latvian: both languages use complementizers originally meaning 'how' – Lithuanian *kaip* and Latvian *kā* – in complements of verbs like 'see', 'hear', and also the verbs of intentional perception, Lith. *žiūrėti* and Latv. *skatīties* 'look', and Lith. *klausyti* and Latv. *klausīt* 'listen'.

Lithuanian

(18) *Kas* *matė* [***kaip*** *pėsči-ąjį* *kliudė* *dviratis*]*?*
who.NOM see.PST.3 **COMP** pedestrian-ACC.SG catch.PST.3 bicycle-NOM.SG
'Who saw a bicycle catching a pedestrian?'
(http://etaplius.lt/kas-mate-kaip-pesciaji-kliude-dviratis/, accessed Sept. 2012)

Latvian

(19) *Zirg-u* *jūg-dam-s,* *Oskar-s* *skatījās,* [***kā*** *māsa*
horse-ACC.SG harness-CVB-M.SG Oscar-NOM look.PST.3 **COMP** sister-NOM.SG
smēla *ūden-i*].
draw.PST.3 water-ACC
'While he harnessed his horse he looked at his sister drawing water.'
(Vilis Lācis, 1904–1966)

3.2 Predicates of fearing

The complementation of verbs of fearing displays an oscillation which shows it to be transitional between propositional and state-of-affairs complementation. This duality has already been observed by Lichtenberk (1995). A subject expressing his apprehension refers to the likelihood that a certain undesirable event may occur (an element shared with epistemic modality) and at the same time expresses his or her wish that it may not occur (an element of volitional, that is, deontic modality). This is reflected, on the one hand, in the alternative use of two complementation strategies, one of the propositional and one of the state-of-affairs type, and, on the other hand, in constructions combining elements of both strategies.

Lithuanian has two types of clausal complements with verbs of fearing like *bijoti* 'fear' and *būgštauti* 'apprehend', but both have the same complementizer *kad*, which is used both for propositional and state-of-affairs complements. The difference is in the use of moods and negation: the propositional type has a realis form of the verb (often a future or a verb combined with the modal verb *galėti* 'can'), whereas the state-of-affairs type has an irrealis form and a negation (the elements combined in apprehensional meaning being 'X fears Y may happen = X thinks Y may happen and wants Y not to happen').

Lithuanian

(20) *Bijau,* [***kad*** *pradės* *lyti*].
fear.PRS.1SG **COMP** start.FUT.3 rain.INF
'I'm afraid it's going to rain.'

Lithuanian

(21) *Bijau,* [***kad*** *ne-pradėtų* *lyti*].
fear.PRS.1SG **COMP** NEG-start.IRR.3 rain.INF
'I'm afraid it might start raining.'

Latvian has two different constructions as well with verbs like *baidīties* 'fear' and *bažīties* 'apprehend'. From what was said above on the twofold nature of complementation with verbs of apprehension one would expect alternative use of the propositional complementizer *ka* and the state-of-affairs complementizer *lai*. In fact the use of the latter is possible but quite rare: usually *ka* is used in both types of complements. Still, they are quite distinct: the state-of-affairs type has the characteristic negation and, instead of the future tense normally used in the propositional type, it has the rigid present-tense form characteristic of state-of-affairs complements in Latvian (also illustrated in (10b) above).

Latvian

(22) a. *Baidos,* [***ka*** *sāks* *līt*].
fear.PRS.1SG **COMP** start.FUT.3 rain.INF

Latvian

(22) b. *Baidos,* [***ka*** *ne-sāk* *līt*].
fear.PRS.1SG **COMP** NEG-start.PRS.3 rain.INF
'I'm afraid it may start raining.'

This latter type is formally mixed: it combines the propositional complementizer *ka* with certain elements of the state-of-affairs type of clausal complementation, especially the temporal rigidity of the verb form, and the negation.

In both Lithuanian and Latvian, verbs of fearing may furthermore combine with the complementizers for interrogation and suspension of judgment, Lith. *ar* and Latv. *vai*. With these complementizers, a complement clause without negation expresses what the subject regards as desirable and fears might be prevented from occurring:

Lithuanian

(23) *Tačiau* *iš* *pradži-ų* *bijojo,* [***ar*** *jos* *parod-a*
however from beginning-GEN.PL fear.PST.3 **COMP** 3.GEN.SG.F exhibition-NOM.SG
sulauks *dėmesi-o,* *ar* *nors* *vien-as* *žmog-us*
get.FUT.3 attention-GEN.SG COMP as.much.as one-NOM.SG.M person-NOM.SG
išdrįs *parašyti* *laišk-ą*].
dare.FUT.3 write.INF letter-ACC.SG
'At first, however, she was afraid whether her exhibition would draw attention, and whether as much as one person would dare to write her a letter.'
(http://www.sekunde.lt/kultura/dailininke-primine-laisku-rasymo-malonuma/, accessed Sept. 2012)

Latvian

(24) *Viņ-š* *paskatījās* *uz* *vis-ām* *pus-ēm,* *kā* *baidī-damies,*
3-NOM.SG.M look.PST.3 to all-DAT.PL.F side-DAT.PL as fear-CVB.NOM.SG.M
[***vai*** *kād-s* *to* *nenoskata*], *un* *paraustīja*
COMP somebody-NOM 3-ACC.SG.PL observe.PRS.3 and pull.PST.3
vārtiņ-u *kliņķ-i.*
back_gate-GEN.PL knob-ACC
'He looked about to all sides as if he was afraid somebody could observe him, and gave the gate knob a jerk.'
(Kārlis Zariņš, 1889–1978)

It is clear that apprehensional verbs are here being drawn into the sphere of epistemic uncertainty, which also comprises a verb like 'doubt'. Once again, it is an illustration of the oscillation of predicates of fearing between propositional and state-of-affairs types of complementation.

3.3 Non-factive marking with commentative predicates

Like many other languages, Lithuanian and Latvian use an adverbial strategy involving conditional or temporal subordinators to cancel the default factive reading of clausal complements with commentative predicates like 'be glad', 'regret', 'like', 'hate', etc. The use of these predicates often presupposes the realization of the relevant state-of-affairs, which accounts for the alternative term 'factive predicates', going back to Kiparsky & Kiparsky (1970). The typically factive character of these predicates is reflected in the use of propositional rather than state-of-affairs complementizers. It is, however, possible to evaluate a certain type of behaviour in an abstract way, without reference to an actual occurrence, and this is often done by expressing the complement in the form of a conditional or temporal clause. The following Latvian examples illustrate the use of the propositional complementizer *ka* (*lai* is never used here) in (32) as well as the adverbial strategy with the conditional subordinator *ja* 'if' in (33):

Latvian

(25) a. *Es* *ne-ciešu,* [***ka*** *man* *runā* *pretī*].
1SG.NOM NEG-suffer.PRS.1SG **COMP** 1SG.DAT speak.PRS.3 against
'I cannot stand being contradicted.'
(Mārtiņš Zīverts, 1903–2000)

Latvian

(25) b. *Viņ-š* *nekad* *ne-cieta,* [***ja*** *kāds* *tam*
3-NOM.SG.M never NEG-suffer.PST.3 **if** somebody 3.DAT.SG.M
mēģināja *izteikt* *līdzjūtīb-u*].
attempt.PST.3 express.INF sympathy-ACC
'He couldn't stand people attempting to express their sympathy for him.'
(Pāvils Rozītis, 1889–1937)

3.4 Irrealis complementizers

In the state-of-affairs domain, Latvian has, in addition to the basic state-of-affairs complementizer *lai*, another complementizer, *kaut*, which is often used if the object of an act of volition is viewed as unreal. This can be seen in the following pair of examples, where there is, apart from the difference in complementizers, an opposition in mood: the irrealis complementizer is used only with irrealis mood, and the low expectation of actuation is additionally emphasized by irrealis marking on the matrix verb. Note that the use of *kaut* is optional and *lai* could also be used:

Latvian

(26) a. *Es* *gribu,* [***lai*** *tu* *klausi* *man-iem*
1SG.NOM want.PRS.1SG **COMP** 2SG.NOM listen.PRS.2SG my-DAT.PL.M
padom-iem].
advice-DAT.PL
'I want you to listen to my advice.'

Latvian

(26) b. *Es* *vēlētos,* [***kaut*** *tu* *klausītu* *man-iem* *padom-iem*].
1SG.NOM wish.IRR **COMP** 2SG.NOM listen.IRR my-DAT.PL.M advice-DAT.PL
'I wish you would listen to my advice.'

The irrealis form could also optionally be used with *lai* in order to mark 'low expectation of actuation', but is obligatory with *kaut*, which apparently functions as an inherently irrealis-marked complementizer. I will return to the notion of *irrealis complementizer* in Section 8, as the term has also been used in a different sense in the literature. At this stage I will define an irrealis complementizer as a complementizer that marks the content of the clausal complement as unreal (either displaying 'low expectation of actuation' or counterfactive).

3.5 Epistemically marked complementizers

A subtype of propositional complementizers is truth-qualifying, or epistemically marked, complementizers. Noonan (2007) mentions languages with complementizers encoding the speaker's assessment of the truth or falsehood of the embedded propositions. He cites the example of the Jacaltec complementizer *tato*, which is "used with complements about which there is some reservation on the part of the speaker, or even outright disbelief" (Noonan 2007: 58–59. It is contrasted with the complementizer *chubil*, which does not convey any qualification.

Similar facts can be cited from Baltic, where the basic propositional complementizers Lith. *kad* and Latv. *ka* can be contrasted with more peripheral complementizers of various origin that shed some doubt on the degree of certainty of the propositional content of the clausal complement. There seem to be two degrees of such epistemic qualification: (i) the content of the complement may be viewed as doubtful, not guaranteed, or (ii) it may be qualified as being decidedly at variance with actual truth. These two degrees are associated with different types of complement-taking predicates.

3.5.1 From comparison to epistemic marking

One source of epistemically marked complementizers is counterfactive comparative subordinators of the type 'as if'. In Lithuanian these can be illustrated by *tarsi*, *tartum*, *lyg*, all of which mean 'as if, as though' and are used in counterfactive comparative clauses:

Lithuanian

(27) *Nors ir* *ne-žino-dam-a* *savo* *ligoni-o* *aukšt-o* *rang-o,*
though NEG-know-CVB-F own patient-GEN.SG high-GEN.SG.M rank-GEN.SG
ji *taip* *rūpestingai* *jį* *prižiūrėjo* *ir*
3.NOM.SG.F so carefully 3.ACC.SG.M look.after.PST.3 and
globojo, [***tarsi*** *jis* *būtų* *jos* *tėv-as*].
nurse.PST.3 **as.if** 3.NOM.SG.M be.IRR.3 3.GEN.SG.F father-NOM
'Though unaware of the high rank of her patient, she looked after him and nursed him as carefully as if he had been her father.'
(www.lcn.lt/bzinios/bz0501/501hom1.html, accessed Sept. 2012)

This strategy can be used in order to mitigate the content of a clausal complement with verbs like *atrodyti* 'to seem': the implication of the use of *tarsi* is that the formulation is actually too strong and the content of the clausal complement needs some qualification.

Lithuanian

(28) *Man* *atrodo,* [***tarsi*** *reiktų* *papildom-ą* *aukšt-ąjį*
1SG.DAT seem.PRS.3 **as.if** be.needed.IRR.3 additional-ACC.SG high-ACC.SG.M.DEF
moksl-ą *baigti,* *kad* *suprastum* *jų*
education-ACC.SG complete.INF in.order.to understand.IRR.2SG 3.GEN.PL
sistem-ą].
system-ACC.SG
'It almost seems to me you would have to get a second university diploma in order to understand their system.'
(http://www.15min.lt/naujiena/zmones/lietuviai/tadas-vidmantas-renka-lesas-komedijai-apie-vonioje-uzsitrenkusia-mergina-3-230925, accessed Sept. 2012)

This is an adverbial strategy of complementation which makes use of a shift from state-of-affairs to propositional clause linking. The typical coincidence of two situation types is reinterpreted in terms of an epistemic evaluation of a proposition: 'I feel as I would have felt if it was getting cold' → 'I feel it's getting cold'. A pathway from similitude to evidentiality ('this sounds like a flute' → 'this very likely is a flute') is pointed out by Fortescue (2010: 139), but it could also be said to lead to epistemic modality (after all, Fortescue has in mind inferentiality, which is very close to epistemic modality), and it is obviously with epistemic rather than evidential marking that we are dealing in the instances discussed here.

But 'seem' is a verb that in itself encodes a weak degree of conviction, so that even combined with a 1st person pronoun, as in *man atrodo* 'it seems to me', there is no incompatibility with the implication of epistemic uncertainty contained in the use of *tarsi* and *lyg* 'as if'. There would obviously be incompatibility in the case of verbs that are inherently associated with a greater degree of epistemic certainty, like *teigti* 'affirm'. Such verbs can, however, be combined with complementizers like *tarsi* and *lyg* when they are used in the 3rd person and refer to other people's assertions. In that case, the function of the complementizer seems to be to contrast other people's claims with the speaker's convictions, which results in an unambiguously negative epistemic evaluation, the content of the clausal complement being qualified as being at variance with reality: 'they raise claims as if they were innocent' → 'they claim to be innocent, though they are not'.

Lithuanian

(29) *antr-oje* ***to*** *sakini-o* *dal-yje*
second-LOC.SG.F **that.GEN.SG.M** sentence-GEN.SG part-LOC.SG
jau *teigia,* [*lyg* *ta* *„Almach“* *iš tikrųjų*
already claim.PRS.3 as.if that.NOM.SG.F Almach actually
darė *poveik-į* *Prezident-ui* *ir* *patarėj-ams*].
make.PST.3 influence-ACC.SG President-DAT.SG and adviser-DAT.PL
'In the next part of that sentence he already claims that this Almach actually exerted influence on the President and his advisers.'
(http://www.mediabv.lt/resursai/publicistika/Politines_technologijos/Paksas/Samokslo_kronika_2003_2004.pdf, accessed Sept. 2012)

In this example we have a realis form of the verb, whereas an irrealis form would be required in the source construction, the counterfactive comparative construction.

In Latvian the same process has affected the compound counterfactive comparative subordinator *it kā*. In (39) it is an adverbial subordinator:

Latvian

(30) *Pēc* *tād-a* *arī* *izskatījās* *un* *arī*
after such-GEN.SG.M also look.PST.3 and also
izturējās [***it kā*** *tam* *vis-s* *piederētu*].
behave.PST.3 **as.if** 3.DAT.SG.M everything-NOM.SG.M belong.IRR
'He looked like one (a rich man) as well and also behaved as if everything belonged to him.'
(http://z6.invisionfree.com/CCLV/index.php?s=13471bb26582db9b93f18e0f8e09346a&showtopic=5037&st=15&#last, accessed Sept. 2012)

But with verbs like *likties* 'seem' it reflects lack of conviction:

Latvian

(31) *Man* *liekas,* [***it kā*** *šodien* *būtu* *vēs-āk-s* *nekā*
1SG.DAT seem.PRS.3 **as.if** today be.IRR chilly-CMP-NOM.SG.M than
vakar].
yesterday
'It seems to me that it's chillier today than yesterday.'(Jānis Parapuķe, 1864–1902)

Finally, with verbs of propositional stance used in the 3rd person and reflecting views held by people other than the speaker, it indicates that the content of the clausal complement is, in the speaker's view, at variance with reality:

Latvian

(32) *Viņ-i* *ne-izprot,* ***ka*** *reform-as* *ir* *viņ-u*
3-NOM.PL.M NEG-understand.PRS.3 **COMP** reform-NOM.PL be.PRS.3 3-GEN.PL
paš-u *lab-ā,* *bet* *domā,* [*it kā* *tā* *būtu*
self-GEN.PL good-LOC.SG but think.PRS.3 as.if that.NOM.SG.F be.IRR
kriev-u *pārlatviskošan-a*].
Russian-GEN.PL Latvianization-NOM.SG
'They do not realize that the reforms are in their own interest but think they are (an attempt at) Latvianising the Russians.'
(http://www.diena.lv/arhivs/ko-doma-krievi-kas-macas-latviesu-skolas-11922467, accessed Sept. 2012)

In the Latvian example (32) we have an irrealis (conditional) form, but this is by no means necessary, as shown by (33):

Latvian

(33) *Šķiet,* *pēc* *katr-as* *stipr-āk-as*
seem.PRS.3 after every-GEN.SG.F strong-CMP-GEN.SG.F
viesuļvētr-as *parādās* *'ekspert-i',* *kas* *apgalvo,* [***itkā***
gale-GEN.SG appear.PRS.3 expert-NOM.PL REL.NOM claim.PRS.3 **as.if**
to *izraisījusi* *globāl-ā* *sasilšan-a*]...
3.ACC.SG cause.PPA.NOM.SG.F global-NOM.SG.F.DEF warming-NOM.SG
'It seems that after every strongish gale 'experts' are popping up who claim that it was caused by global warming...'
(http://klab.lv/~ritvars/, accessed Sept. 2012)

The shift from irrealis to realis marking shows that we have a shift from irrealis marking (mentioned in Section 3.4), i.e., a strategy for marking the content of a clause as unreal, to a type of epistemic marking, i.e., a means of expressing doubt or scepticism as to the content of the clausal complement. Whereas (37) and (40) are intermediate between the two (low degree of confidence is reflected in the use of irrealis marking), in (38) and (42) irrealis marking in the verb is absent and the counterfactive subordinators have shifted to the function of epistemically marked complementizers. I will return to the distinction in Section 8.

3.5.2 From hearsay to epistemic marking

A second type of development is illustrated only in Lithuanian. It concerns the evidentiality marker *esą*, which is the outcome of a degrammaticalization process: it has been extracted from evidential forms of the verb. In Lithuanian, the evidential is derived by replacing a finite verb form with a participle; if the

verb form is a compound, and consists of an active past participle and a form of the verb *būti* used as an auxiliary, it is the auxiliary that will receive the evidential marking and correspondingly appear in the participial form:

Lithuanian

(34) *Jis* *es-ąs* *išvažiav-ęs*.
3.NOM.SG.M be-PRPA.NOM.SG.M depart-PPA.NOM.SG.M
'He is said to have left.'

Such constructions as in (34) are, however, becoming increasingly rare: instead of the participial form of 'be', which would have to agree with the clausal subject, we now have an undeclinable form *esą*, originally a neuter singular form. It functions as an evidentiality marker and is no longer part of a compound verb form: the main verb can still be in the participial form, but a finite form is possible as well:

Lithuanian

(35) a. *Jis* *esą* *išvažiavęs*.
3.NOM.SG.M EVID depart.PPA.NOM.SG.M

Lithuanian

(35) b. *Jis* *esą* *išvažiavo*.
3.NOM.SG.M EVID depart.PST.3
'He allegedly left.'

This new adverbial evidentiality marker seems to be exclusively a hearsay marker, which is a narrowing of the original meaning, as the Lithuanian verbal evidential can be reportive, inferential or mirative (Holvoet 2007: 84–85). We see this hearsay marker appear not only in simple sentences, as in (43), but also in complement clauses with verbs of saying:

Lithuanian

(36) *Per* *apklaus-ą* *teisėj-a* *sakė,* [***kad*** *esą* *paraš-as*
during hearing-ACC.SG judge-NOM.SG say.PST.3 **COMP** esą signature-NOM.SG
jos, *bet* *kaip* *jis* *atsirado,* *pasakyti* *ne-gali*].
3.GEN.SG.F but how 3.NOM.SG.M get.PST.3 say.INF NEG-be.able.PRS.3
'During the hearing the judge said that {esą} the signature was hers but she couldn't tell how it got there.'
(http://kauno.diena.lt/naujienos/kaunas/miesto-pulsas/dmeksraitienes-advokatei-nepavyko-nusalinti-teisejo-331274, accessed Sept. 2012)

It is not quite clear what kind of function *esą* performs in such cases. A hearsay marker is not strictly needed here in view of the presence of a verb of saying, but one could imagine *esą* functioning as a redundant evidential marker. It seems, however, that the very redundancy of this hearsay marker in the clausal complement should favour an epistemic reinterpretation.

Often the complementizer *kad* is dropped and *esą* appears to have taken over the function of complementizer. The question could be raised whether in such instances *esą* has really become a complementizer or whether these are instances of complementizer deletion. As mentioned in Section 5, complementizer deletion seems to occur in spoken language, but there are cases where it is clearly impossible, and one of these cases is complement clauses of nouns, which are, of course, characteristic of a more formal language, where complementizer deletion does not occur:

Lithuanian

(37) [...] *mint-is,* [***esą*** *universitet-ai* *turi* *gaminti*
thought-NOM.SG **COMP** university-NOM.PL should.PRS.3 produce.INF
darbuotoj-us *ir* *tik* *tuos,* *kuri-ų* *reikia*
worker-ACC.PL and only that.ACC.PL REL-GEN.PL be.needed.PRS.3
darb-o *rink-ai*], *yra* *kritikuotin-a*.
labour-GEN market-DAT.SG be.PRS.3 open.to.critisism-NOM.SG.F
'The notion {esą} the universities should produce workers, and only such as are needed on the labour market, is open to criticism.'
(http://www.vu.lt/site_files/SD/Naujienlaiskis/2012_Stojantiems_i_magistrantura.pdf, accessed Sept. 2012)

What exactly is the function of the complementizer *esą*, if a complementizer it is indeed? If we assume a connection with the original hearsay function of the adverb *esą*, we could venture that *esą* is a kind of evidential complementizer; to be more precise, it would be a reportive complementizer, since *esą* does not seem to have other functions than hearsay. What could be the *raison d'être* of a reportive complementizer? To the extent that it would occur with verbs of saying it would be redundant, which, of course, does not mean that it cannot exist. Evidential (hearsay) marking on the verb is frequent in complements of verbs of saying in Latvian, for instance, though it is redundant here as well. It is shown in the following example, where the complement clause verb is in the oblique (evidential) mood:

Latvian

(38) *Viņ-š* *teica,* [***ka*** *es-ot* *aizņemt-s*].
3-NOM.SG.M say.PST.3 **COMP** be-OBM busy-NOM.SG.M
'He said he was busy.'

We could therefore also expect redundant marking in the complementizer. But in a sentence like (47), or with verbs of propositional stance in general, there would be no redundancy. It is conceivable that *mintis, esą* means something like 'the notion that, as some people say...'. But if *esą* is evidential in (47), it also has strong epistemic overtones. There must be a discrepancy between what is reported or claimed and what the speaker himself thinks. As soon as we change the assessment of the idea referred to in (47) from negative to positive, the sentence with *esą* becomes deviant:

Lithuanian

(39) **[...] *mint-is,* [***esą*** *universitet-ai* *turi* *gamin-ti*
thought-NOM.SG **COMP** university-NOM.PL should.PRS.3 produce-INF
darbuotoj-us *ir* *tik* *tuos,* *kuri-ų* *reikia*
worker-ACC.PL and only that.ACC.PL.M REL-GEN.PL be.needed.PRS.3
darbo *rink-ai*], *yra* *teisinga-a*.
labour-GEN.SG market-DAT.PL be.PRS.3 correct-NOM.SG.F
Intended meaning: 'The notion that the universities should produce workers, and only such as are needed on the labour market, is correct.'

This suggests that there must be a negative epistemic evaluation of the contents of the clausal complement in order for *esą* to be used. The mechanism of the rise of this epistemic meaning is, of course, closely linked with reportive meaning: there is a discrepancy between what is reported or claimed and what the speaker himself thinks, and this discrepancy can only lead to negative, not to positive epistemic evaluation. Examples like (46) show, however, that this epistemic evaluation can become dominant: it becomes irrelevant whether anybody has actually advanced the claim contained in the clausal complement, and *esą* comes to mean something like 'as one might be tempted to think'.

3.5.3 Epistemically marked complementizers with pretence predicates

The two strategies outlined in the two preceding sections are also applied in Lithuanian and Latvian to pretence predicates. Noonan notes that the complements of such predicates are consistently treated as realis (indicative) in languages with a mood distinction. As the putative reason for this he opines that pretence predi-

cates "establish an alternative reality and the complement constitutes an assertion within that alternative reality" (Noonan 2007: 126–127). The Baltic facts seem to confirm this surmise. The basic type of complementation here is the propositional one, with Lith. *kad* and Latv. *ka*, both used with the realis (indicative):

Lithuanian

(40) *Britanij-os* *valdži-a* *vis dar* *apsimeta,* [*kad* L
Britain-GEN authorities-NOM.SG still pretend.PRS.3 COMP
kontroliuoja *imigracij-ą*].
control.PRS.3 immigration-ACC
'The British authorities keep pretending they control immigration.'
(http://www.londonozinios.com/v2/britanija/aktualijos-d-britanija/7928-britanijos-valdzia-vis-dar-apsimeta-kad-kontroliuoja-imigrac.html, accessed Sept. 2012)

Latvian

(41) *Vien-i* *izliekas,* [***ka*** *maksā* *nodokļ-us*], *otr-i* –
one-NOM.PL.M pretend.PRS.3 **COMP** pay.PRS.3 tax-ACC.PL other-NOM.PL.M
[*ka* *tos* *iekasē*].
COMP 3.ACC.PL.M collect.PRS.3
'Some pretend they pay taxes, others pretend they collect them.'
(http://twitter.com/KristapsMors/status/9277231831, accessed Sept. 2012)

There are, however, alternative strategies. Both Lithuanian and Latvian have the counterfactive comparative subordinator strategy ('pretend, do as if'), here illustrated from Latvian. Note that though the use of irrealis (conditional) marking on the verb would be possible (in fact, it may be inherited from the source construction, counterfactive comparative adverbial clauses), realis forms are found quite frequently, which shows that this is not an irrealis strategy but the use of counterfactive comparative subordinators serves to mark that the content of the clausal complement is at variance with reality, that is, this is a type of epistemic marking rather than irrealis marking.

Latvian

(42) *Sākum-ā* *sieviet-e* *izlikās,* [***it kā*** *viņ-ai*
start-LOC woman-NOM pretend.PST.3 **COMP** 3-DAT.SG.F
ne-sāp] *un* *viņ-a* *nekādi* *ne-reaģēja*.
NEG-hurt.PRS.3 and 3-NOM.SG.F no_way NEG-react.PST.3
'At the start the woman pretended she felt no pain, and she reacted in no way.'
(http://klad.lv/users/zvejnieks/10272.html?mode=reply, accessed Sept. 2012)

Lithuanian has also a second strategy with pretence predicates, namely the use of the complementizer *esą*, which, as noted in the preceding section, originated as an evidential marker:

Lithuanian

(43) *Vis* *apsimeta,* [***esą*** *dirba* *Lietuv-ai*].
always pretend.PRS.3 **COMP** work.PRS.3 Lithuania-DAT
'They keep pretending they work for Lithuania.'
(http://m.lrytas.lt/?id=12749718111274258312&view=6, accessed Sept. 2012)

As it seems, this is not really evidential marking: pretence predicates are not speech act predicates, as pretence involves non-verbal alongside verbal behaviour. True, evidential marking on the verb (oblique mood) is occasionally found in complements of pretence predicates in Latvian, but the unequivocally epistemic function of *esą* as a complementizer noted in the preceding section suggests that here as well, *esą* is a truth-qualifying complementizer, used to mark that the content of the clausal complement is at variance with reality.

3.5.4 Quotative/interpretive complementizers

Two other particles, *atseit* and *neva*, seem to be undergoing the same syntactic development as *esą*. They can occur in independent sentences, but also in embedded clauses. In the latter case they either co-occur with the propositional complementizer *kad*, or they occupy the position of complementizers themselves. These two situations are illustrated in (53) and (54) respectively; they are here shown for *neva*, but the use of *atseit* is similar:

Lithuanian

(44) *Mint-is,* [*kad* ***neva*** *toje* *viet-oje*
thought-NOM.SG COMP **PRT** that.LOC.SG.F place-LOC.SG
galėtų *bū-ti* *kaim-o* *turizm-o* *sodyb-a*] *irgi*
be.able.IRR.3 be-INF country-GEN tourism-GEN farm-NOM.SG also
nuveda *į* *fantazij-ų* *pasaul-į.*
lead.PRS.3 into phantasy-GEN.PL world-ACC.SG
'The notion that {neva} there could be an agritourism farm on this spot leads into the realm of phantasy.'
(http://www.kiaules.lt/index.php?page=1&text=81, accessed Sept. 2012)

Lithuanian

(45) *Dar* *keisčiau* *skamba* *mint-is,* [***neva*** L
still stranger sound.PRS.3 thought-NOM.SG **PRT**
Lietuv-a *tap-usi* *nepriklausom-a*
Lithuania-NOM become-PPA.NOM.SG.F independent-NOM.SG.F
tuomet, *kada* *jos* *Nepriklausomyb-ę* *pripažino*
then when 3.GEN.SG.F independence-ACC recognize.PST.3
užsienio *valstyb-ės*].
foreign states-NOM.PL
'The notion {neva} Lithuania became independent at the moment when other states recognized its independence sounds even more strange.'
(http://www3.lrs.lt/pls/inter/www_tv.show?id=84319,6619,40)

Whereas the etymology of *esą* is transparent, that of *neva* is obscure, so that we cannot reconstruct its original meaning. When used in independent sentences, *neva* seems to convey both evidential (hearsay) and epistemic meanings: the content of the clause is qualified as being based on hearsay or as being doubtful. In complement clauses as illustrated in (54) and (55), *neva* seems to have become basically epistemic: as in the case of *esą*, its use becomes impossible if the element of negative epistemic evaluation is absent.

Apart from the evidential and epistemic functions referred to above, *atseit* and *neva* can have a third function, which we could call *quotative*. It is often used to introduce an utterance reproduced in its original form, i.e., retaining all its original formal features such as 1st and 2nd person pronoun (without person switch required in indirect speech), original discourse markers etc. From quotation markers proper (such as *quote ... unquote* used in journalese) these markers differ in that the utterance they refer to need not be quoted literally, but a free interpretation of it is given. I give two examples, one in which *neva* occurs after the complementizer *kad*, and one where its seems to have taken the position of the complementizer itself.

Lithuanian

(46) *Taip* *kankinomės* *dvi* *savait-es,* *gydytoj-as*
so agonize.PST.1SG two.ACC.PL week-ACC.PL doctor-NOM.SG
sakė, [*kad* ***neva*** *neram-us* *vaik-as,*
say. PST.3 COMP **PRT** nervous-NOM.SG.M child-NOM.SG
nieko *ne-padarysi,* *turėkit* *kantryb-ės*].
nothing NEG-do.FUT.2SG have.IMP.2PL patience-GEN.SG
'So we suffered two weeks of torment and the doctor said that {neva} it's an excitable child and there is nothing to be done, you must be patient.'
(http://www.mamanija.lt/klausimai/15453/mamos-padekite-kas-esat-susidurusios-supanasia-miego-ir-valgymo-problema, accessed Sept. 2012)

Lithuanian

(47) *Jie* *man* *sakė,* [***neva*** *ko* *ne-duodi*
3.NOM.PL.M 1SG.DAT say.PST.3 **neva** why NEG-give.PRS.2.SG
pasimatuoti *šalm-o*], *nors* *kit-as* *jau*
try.on.INF helmet-GEN.SG though other-NOM.SG.M already
buvo *jį* *užsidėj-ęs.*
be.PST.3 3.ACC.SG.M put.on-PPA.NOM.SG.M
'They said to me {neva} why don't you allow us to try on the helmet, though another [of them] had already put it on.'
(http://klaipeda.diena.lt/dienrastis/lietuva/pribloske-ne-tik-museikos-173244#axzz25lO8z81x, accessed Sept. 2012)

It is well known that the borderline between direct and indirect speech is not always strict. Many languages, such as Classical Greek, have complements with verbs of saying introduced by a complementizer but retaining the form of the reproduced utterance, without person switch etc., as illustrated by (58):

Classical Greek

(48) *eĩpon* [***hóti*** *hikan-oí* *esmen*].
say.AOR.3PL **COMP** ready-NOM.PL be.PRS.1PL
'They said they were prepared.'
(Xenophon, cited after Goodwin 1900: 315)

The retention of the original form of the utterance may be licenced by special segmental markers such as *neva* in (57), and such markers may evidently themselves advance to the position of complementizers. We could therefore speak of a special category of quotative complementizers. The strange thing, however, which calls for an explanation, is that *atseit* and *neva* are also used to convey evidential and epistemic meanings; why the same marker should be used to mark quotation and epistemic evaluation is not clear. A possible explanation would be that *atseit* and *neva* originated as undifferentiated evidential and quotative markers, and that the epistemic meaning evolved from the evidential one.

Another explanation involves the notion of interpretive use, introduced by Sperber & Wilson (1986: 228–229). This notion refers to the use of utterances to refer to other utterances rather than to extralinguistic state-of-affairs. The most obvious instance of this is of course quotations, but apart from these interpretive use also comprises what Sperber and Wilson call "echoic interpretations" – these are utterances echoing (through similarity in form) other people's actual utterances or thoughts imputed to them. As Blakemore (2002: 85) notes, languages tend to develop "means for encouraging hearers to interpret utterances as interpretive representations of thoughts attributed either to other speakers or to themselves at other times". Lithuanian *atseit* and *neva* often perform this func-

tion. Whereas hearsay markers typically qualify assertions or at least utterances with propositional content, and their use presupposes an actual verbal utterance (failing which there can be no transfer of information, which is crucial for evidentials), such restrictions do not hold for interpretive use markers. A typical instance is the following, containing the marker *atseit*:

Lithuanian

(49) *Jis* *tik* *papurtė* *galv-ą,* ***atseit*** *ne-trauk,*
3.NOM.SG.M only shake.PST.3 head-ACC.SG **PRT** NEG-pull.IMP.2SG
o *po to* *apsižiūri,* *kad* *dančio* *nėra*.
but after.that realize.PRS.3 COMP tooth-GEN.SG be.PRS.3.NEG
'He only shook his head {atseit} 'Don't pull', but a moment later he realized the tooth was out already.' ('.. as if to say 'Don't pull'')
(http://www.supermama.lt/forumas/lofiversion/index.php/t374-100.html, accessed Sept. 2012)

Though, in the case of interpretive use markers, assertion is not required, these markers will acquire additional meaning shades in those cases where assertion is involved. Whereas quotations proper can be perfectly stance-neutral, echoic interpretations usually are not. As Sperber & Wilson note, "these interpretations achieve relevance by informing the hearer of the fact that the speaker has in mind what so-and-so said, and has a certain attitude to it" (Sperber & Wilson 1986: 238). This attitude is normally negative. Whereas in the case of evidentials acquiring epistemic overtones the elements of negative evaluation concern the truth or falsehood of the evidentially marked information, interpretive use markers may be used in imputing an utterance or thought characterising a person's mentality or non-verbal behaviour. These two elements cannot always be clearly set apart: in (56), for example, we do not know whether *neva* is used to characterize the doctor's tone and behaviour or the truth of what he/she says.

There seems, at any rate, to be a domain of discourse-relevant linguistic marking that comprises epistemic, evidential and interpretive use marking, typically expressed by adverbs/particles, but capable of spreading to complementizer position – in agreement with Frajzyngier's idea (Frajzyngier 1995: 475) that complementizers are one of the syntactic slots in which modal meanings – but we could add evidential meanings and interpretive use marking – may be expressed.

The Lithuanian developments outlined here are, at any rate, still ongoing and the complementizer status of the markers discussed in this section is still open to controversy. This question deserves to be discussed in a broader context. It is generally known that verbs of saying, as well as those denoting similarity, give rise to complementizers (Heine & Kuteva 2002: 257–258, 261–265, 273–275). At some stage they must be quotative in the sense of retaining, in their comple-

ments, the original person and tense forms as in quoted direct speech. Noonan (2007: 122) considers this to be evidence against the interpretation of a marker containing the verb *bɦi-* 'say' as a complementizer in Chantyal:

Chantyal

(50) [*na* *tisuŋ* *Kadmandu-ri* *ɦya-i* *bɦi-wa*]
I last.year Kathmandu-LOC go-PFV say-NMLZ
khi-sə *səmjɦi-i*
he-ERG remember-PFV
'He remembered that he went to Kathmandu last year'
(literally: 'He remembered saying, "I went to Kathmandu last year."').

As the sentence actually means 'he remembers going to Kathmandu' rather than 'he remembers saying he went to Kathmandu', grammaticalization already seems to have occurred, and one wonders why *bɦi-wa* should not be a complementizer. Noonan does not argue this point, but perhaps it rests on the intuition that a quotation is an alien body within the syntactic context in which it occurs, and that between the marker introducing it and the quotation itself there cannot be a regular relation as between complementizer and complement. This may be so, but quotations proper are only one form of interpretive use in the sense of Sperber & Wilson (1986). Echoic interpretations are considerably better integrated in the syntactic context, and between complement clauses preserving part of the formal features of the replicated utterance (person, tense) and other types there is no very sharp line of division. Echoic elements can often be found even in complement clauses with otherwise well-established complementizers like *kad*. Take the following example, where an imperative is used in a complement clause with a speech act verb, actually with the value of a deontic modal predicate rather than of a directive:

Lithuanian

(51) *Pavyzd-ys* *apie* *nam-ą,* *kur-į* *nugriauti* *liepė*
example-NOM.SG about house-ACC REL-ACC.SG.M pull.down.INF order.PST.3
valdži-a, *sako,* [***kad*** *dirbk* *taip,* ***kad*** *niekas*
authorities-NOM say.PRS.3 **COMP** work.IMP.2SG so **COMP** nobody.NOM
ne-pamatytų].
NEG-see.IRR.3
'The example of the house which the authorities ordered to be pulled down tells us that one should work without being seen by anybody' (literally: '... that work (imperative!) without being seen by anybody').
(http://www.delfi.lt/news/ringas/lit/rsimasius-reguliavima-perki-absurdas-nemokamai.d?Id=16016282, accessed Sept. 2012)

More or less the same situation can be observed in Latvian. The following example contains a 3rd person hortative rather than an imperative, but this hortative construction is deviant as well in a complement clause with *ka*, where one would normally expect a formally assertive, propositional construction:

Latvian

(52) *Kraukl-is* *jau* *apgalvo,* [***ka*** *Caun-i*
Krauklis-NOM already claim.PRS.3 **COMP** Caune-ACC
lai *gaid-ot* *tikai* *pēc* *gad-iem*].
HORT expect-OBM only after year-DAT.PL
'Krauklis is already claiming that Caune is not to be expected back for years.'
literally '... is claiming that let them not expect Caune back for years.'
(Pāvils Rozītis, 1889–1937)

Such examples are, admittedly, on the periphery of complementation strategies, and many languages do not tolerate them at all. However, they may achieve quite a high degree of grammaticalization in other languages, and special quotative and/or interpretive complementizers may develop; this is a development parallel to that of what I have called *interpretive deontics*, cf. Holvoet and Konickaja (2011).

4 On the history of Baltic complementizers

Old Lithuanian had a distinction between propositional and state-of-affairs complements equivalent to modern Latvian *ka* and *lai*. The now universal *kad* was once restricted to state-of-affairs complements whereas *jog* was used for propositional complements. Some Old Lithuanian authors use *idant* rather than *kad*, but observe the same distinction.

Old Lithuanian

(53) *Kas* *pasakie* *tau* [***jog*** *esi* *nogu*]*?*
who tell.PST.3 2SG.DAT **COMP** be.PRS.2SG naked.INS.SG.M
'Who told thee that thou wast naked?'
(Chyliński Bible, 17th century, Genesis 3:11, cited from Kavaliūnaitė, ed., 2008)

Old Lithuanian

(54) *Walgieygu* *isz ano medzia,* *ape kuri* *prysakiau*
eat.Q from that tree about which command.PST.1SG
tau [***kad*** *newalgitumey* *isz jo*]*?*
2SG.DAT **COMP** NEG.eat.IRR from it
'Hast thou eaten of the tree, whereof I commanded thee that thou shouldest not eat?'
(Chyliński Bible, 17th century, Genesis 3:11)

The age of this distinction is not known, but its origins can be surmised. *Jog* is derived from the anaphoric-relative pronominal stem **yo-* and actually seems to have residual relative functions in Old Lithuanian texts:

Old Lithuanian

(55) *Kas* *tey* *ira* [***jog*** *padar-ey*]*?*
what.NOM that be.PRS.3 **COMP** do-PST.2SG
'What is this that thou hast done?'
(Chyliński Bible, 17th century, Genesis. 3:13)

When we look at adverbial clauses, we see that *jog* is also used in consecutive clauses introduced by *taip* 'so':

Old Lithuanian

(56) *Ir* *twino* *wądeniey,* *ir* *pakiełe* *Karobli,*
and increase.PST.3 water.NOM.PL and lift.PST.3 ship.ACC
teyp [***jog*** *pasikiełe* *wirszuń* *źiames*].
so **COMP** rise.PST.3 above earth.GEN
'And the waters increased, and bare up the ark, and it was (so that it was) lift up above the earth.'
(Chyliński Bible, 17th century, Genesis 7:17)

Kad is derived from *kada*, which still exists in the meaning 'when' (pronominal adverb or adverbial subordinator), cf. Smoczyński (2007: 242). *Kad* is, however, used in several types of adverbial clauses, including purpose clauses:

Old Lithuanian

(57) *uźstate* *ghi* *Sode* *Eden,* [*kad*
put.PST.3 3.ACC.SG.M garden.LOC Eden SUB
apdirbtu *ghi,* *ir* *saugotu* *jo*].
dress.IRR.3 3.ACC.SG.M and keep.IRR.3 3.GEN.SG.M
'and put him into the garden of Eden to dress it and to keep it'
(Chyliński Bible, 17th century, Genesis 2:15)

This purposive construction is formally nearest to constructions with the complementizer *kad*, also with regard to mood, which is always irrealis in both instances in Old Lithuanian. It is thus conceivable that the Old Lithuanian distinction between a propositional and a state-of-affairs complementizer reflects two complementation strategies mentioned by Dixon (2006: 6), namely a purposive strategy (*kad*) and a relative strategy (*jog*).

The Latvian distinction between *ka* and *lai* is an internal Latvian development. *Lai* comes from **laid*, the imperative of *laist* 'let', and originated as a hortative particle in sentences like

Latvian

(58) *Lai* *viņš* *atnāk.* L
HORT 3.NOM.SG.M come.PRS.3
'Let him come.'

Subsequently *lai* (from an original form *laid*, the 2nd singular imperative of *laist* 'let') spread to complements of verbs like 'order', 'want' etc. as well as to several types of adverbial clauses, such as purpose clauses.

Latvian *ka* is assumed to be from Indo-European **k*^{w}*od*, the neuter form of the interrogative relative pronoun 'what' (Endzelin 1923: 398). If this derivation is correct, *ka* is not etymologically identical with Lithuanian *kad*, and none of the complementizers now used in the two languages can be traced back to a Common Baltic complementizer. On etymological grounds we may assume that Latvian *ka* continues a relative complementation strategy.

We may also probably assume that before the spread of *lai* the complementizer *ka* combined both functions, that of propositional and state-of-affairs complementizer; there is, in Latvian, nothing similar to the Old Lithuanian propositional complementizer *jog*, so *ka* would functionally more or less correspond to Modern Lithuanian *kad* (propositional and state-of-affairs), though the two are probably not identical etymologically. The Old Latvian texts do not enable us to reconstruct this fragment of the history of complementizers, as their quality is very poor and German influence is very strong. If the assumption formulated here is correct, the use of *ka* in complements of apprehensional predicates might be a trace of an earlier state of the language in which *ka* served both the propositional and the state-of-affairs domains: *ka* would have been replaced with *lai* in state-of-affairs complements with the exception of complements of apprehensional verbs because of the transitional nature of these predicates between epistemic and deontic modality, and thereby also, as far as complementation is concerned, between propositional and state-of-affairs complementation.

The Latvian complementizer *kaut*, the origin of which has long been unclear, has been recently explained by Petit (2012), who posits Indo-European *k*^{w}*ā-ute* 'if however', originally a concessive subordinator, which subsequently acquired a volitional meaning and hence came to be used in complements of verbs of wishing.

In the prehistory of Lithuanian and Latvian, irrealis complementizers of the Slavonic type, incorporating a formal irrealis marker (originally an enclitical form of the verb 'be' used as an irrealis auxiliary) seem to have existed. Their existence in Old Lithuanian was discovered by Stang (1970). The only modern variety of Baltic to have retained traces of this is the Latgalian dialects of Latvian, where the irrealis complementizer *kab* (< *ka* + *bi*) is opposed to the realis complementizer *ka* (Nau 2012: 99):

Latgalian Latvian

(59) *A* *taŭs* *na* *gryb,* [***kab*** *jis* *zogtu*].
but father.NOM.SG NEG want.PRS.3 **COMP** 3.NOM.SG.M steal.IRR
'But his father doesn't want him to steal.'

5 Complementizer omission

Standard Lithuanian and standard Latvian have no complementizer deletion. In the colloquial languages, however, the situation is different and the complementizers Lith. *kad*, Latv. *ka* 'that' are often omitted, at least with speech act verbs. This property of the spoken language is reflected in informal Internet texts. It is here illustrated for Lithuanian (70) and Latvian (71):

Lithuanian

(60) *Su* *daktar-u* *kalbėjau,* *pasakė* [*padarys* *viską,* *kas*
with doctor-INS talk.PST.1SG say.PST.3 do.FUT.3 everything.ACC REL.NOM
įmanom-a, *operacij-ų* *toki-ų* *daro* *ir* *tai* *ne*
possible-N.SG operation-GEN.PL such-GEN.PL do.PRS.3 and that NEG
naujien-a].
novelty-NOM.SG
'I talked to the doctor and he said he would do everything he could, they perform such operations and there is nothing new about that.'
(http://www.supermama.lt/forumas/lofiversion/index.php/t755376.html, accessed Sept. 2012)

Latvian

(61) *Saimniec-ei* *nodevu* *instrukcij-as...* *teica*
mistress-DAT give.PST.1SG instructions-ACC.PL say.PST.3
[*darīs,* *ko* *varēs*].
do.FUT.3 REL.ACC be.able.FUT.3
'I gave instructions to [the dog's] mistress... she said she would do what she could.'
(http://www.suni.lv/viewtopic.php?f=3&t=27&start=260, accessed Sept. 2012)

The conditions and extent of this deletion should be investigated separately. This complementizer deletion should be kept in mind when examining the possible shift of certain adverbs (discourse markers) to the syntactic position of complementizer (cf. Section 3.5.2): it is not always quite clear whether a discourse marker has taken the place of a complementizer or whether one is dealing with complementizer omission.

6 Complementizer combination

Combined complementizers are not a regular feature in Baltic complementation, but combined markers like *kad esą* in (46) and *kad neva* in (56) above suggest that the initial position in complement clauses is one in which two competing needs, namely that of providing formal marking for the syntactic relation of complementation and that of marking 'modality' in a broad sense, may vie for one syntactic slot but may also result in clusterings of complementation markers. The assumption that there is one syntactic slot reserved for a complementizer leads us to claim that *esą* and *neva* are discourse particles as long as the complementizer proper (*kad*) occurs, and to consider them as possible complemetizers only when *kad* is dropped. The linguistic reality is probably more complex.

Of course, what looks like combinations of complementizers may just reflect a transitional stage in the rise of new complementizers. In Latvian, the complementizers *ka* and *lai* are not combined nowadays, but as late as the early 20th century one finds the combination *ka lai*:

Latvian

(62) *Šrāg-as* *un* *nosacījum-i* *prasīja,* [***ka*** *lai*
statute-NOM.PL and regulation-NOM.PL demand.PST.3 **COMP** COMP
ne-vāc-i *no* *vāc-iem* *tiktu* *stingri* *nošķir-t-i*].
neg-German-NOM.PL from German-DAT.PL AUX.IRR strictly segregate-PPP-NOM.PL
'The *Schragen* (guild statutes) and regulations demanded that non-Germans should be strictly segregated from Germans.'(Augusts Deglavs, 1862–1922)

In the modern language, only *lai* would be left. As mentioned in Section 4, *lai* originated as a hortative particle, and constructions like (72) are evidently a trace of an older stage of development. At this transitional stage, tasks seem to be divided: *ka* functions as a syntactic marker of complementation whereas *lai* carries the (obligative) modal meaning.

7 Polyfunctionality of complementizers

7.1 Complementizers and subordinators

The Lithuanian complementizer *kad* is a shortened form of *kada* 'when', and even in modern Lithuanian it often performs the function of temporal or conditional adverbial subordinator alongside that of complementizer.

In Latvian the complementizer *ka* is easier to distinguish from adverbial subordinators, although there is, in colloquial Latvian, a tendency to use the subordinator *kad* 'when' instead of *ka* 'that'. One overlap that is certainly old is the use of *ka* as a subordinator of purpose. This final *ka* occurs almost only with negation and seems to be associated with apprehensional meaning and formally coincides with the complementation type used for verbs of fearing; that the same strategy is used for negative purpose and apprehensional complementation has already been noted (Lichtenberk 1995):

Latvian

(63) *Es* *pat* *baidos* *no* *viņ-iem* *strādniek-u*
1SG.NOM even fear.PRS.1 of 3-DAT.PL.M worker-GEN.PL
māj-ām *skursteņ-us* *mūrēt,* [***ka*** *ne-sabrūk*].
house-DAT.PL chimney-ACC.PL build.INF **SUB** NEG-collapse.PRS.3
'I am even afraid to build chimneys for my workers' houses from them (that is: from these bricks) for fear that they might collapse.' (Pāvils Rozītis, 1889–1937)

The Latvian state-of-affairs complementizer *lai* is also used as an adverbial subordinator in several types of adverbial clauses – purpose, consecutive, concessive. These adverbial clauses are also predominantly of the state-of-affairs type. Both types of use – as adverbial subordinator and complementizer – have evolved from a common source: a hortative particle.

7.2 Complementizers and discourse markers

Both in Lithuanian and in Latvian the interrogative complementizer is also used as an interrogation marker in simple sentences, probably as a result of desubordination:

Lithuanian

(64) ***Ar*** *tu* *žinai,* *kas* *atsitiko?*
COMP 2SG.NOM know.PRS.2SG what.NOM happen.PST.3
'Do you know what happened?'

Latvian

(65) ***Vai*** *tu* *zini,* *ko* *tas* *nozīmē?*
COMP 2SG.NOM know.PRS.2SG what.ACC that mean.PRS.3
'Do you know what that means?'

Kad is frequently used as a discourse marker in spoken Lithuanian. Sawicki (2012) describes it as a responsive particle conveying "various shades of skepticism, reservations or uneasiness of the speaker about the content of the previous turn and offer justifications or excuses for not complying or obeying (Sawicki 2012: 151)".

Lithuanian

(66) *Supranti,* *ką* *aš* *noriu* *dabar* L
understand.PRS.2SG what.ACC 1SG.NOM want.PRS.1SG now
pasakyti? – *Kad* *gal* *ne-labai,* *dėd-e.*
say.INF COMP maybe not-much uncle-VOC
'Do you understand what I want to tell you now? – Well maybe not really, uncle.'

Whether this *kad* is indeed the complementizer or whether it originated as an adverbial subordinator is hard to establish; Sawicki (2012) notes that this use corresponds to that of Polish *kiedy*, which is an adverbial subordinator meaning 'when' – a sense that is also attested for Lithuanian *kad*, though it is now obsolete in the standard language.

Another use of *kad* that might be related to the complementizer use is in wishes:

Lithuanian

(67) *Kad* *aš* *būčiau* *toks* *išminting-as* *kaip* *tu!* L
COMP 1SG.NOM be.IRR.1SG such wise-NOM.SG.M as 2SG.NOM
'If only I were as wise as you!'

Again, though it is conceivable that this use is to be connected with that of *kad* as a complementizer, adverbial origin cannot be excluded, as *kad* has conditional uses as well, illustrated in

Lithuanian

(68) *Kad* *žinočiau,* *pasakyčiau.*
if know.IRR.1SG say.IRR.1SG
'If I knew, I would tell.'

Counterfactive conditional markers being a common source of wish markers (cf. English *if only* etc.), the exact nature of this optative *kad* remains unclear: it could be the complementizer or the conditional subordinator.

In Latvian the complementizer *ka*, used with apprehensional verbs, is also used in simple apprehensional sentences:

Latvian

(69) *Ka* *tik* *tante* *ne-pamana!*
COMP only aunt NEG-notice.PRS.3
'If only auntie doesn't notice!' (Jānis Greste, 1876–1951)

It is not completely clear that this *ka* is indeed the complementizer, because *ka* is, as mentioned above, also used in a certain type of final clauses, and, as Lichtenberk (1995) notes, negative final clauses (*lest*-clauses) are a frequent source of apprehensional markers across clause types.

8 More on semantic distinctions in complementizers

In addition to the widely accepted distinction between propositional and state-of-affairs complementizers, two other semantic distinctions have been invoked in this article: realis vs irrealis and epistemically neutral vs epistemically marked complementizers.

The term *irrealis complementizer* has been used in the literature to refer to state-of-affairs complementizers (Amman & Van der Auwera 2004). However, the Latvian opposition between the state-of-affairs complementizer *lai* and the additionally irrealis-marked *kaut* shows that though the two notions may coincide, irrealis complementizers can also be viewed as a subtype within state-of-affairs complementizers. Secondly, the realis-irrealis distinction can also encroach upon the domain of propositional complementizers. This may be seen in Slavonic languages such as Russian and Polish, which have developed complementizers comprising formal irrealis markers. Cf. Russian:

Russian

(70) *On-a* *chočet,* [***čtoby*** *ja* *ej* *pomog*].
3-NOM.SG.F want.PRS.3SG **COMP** 1SG.NOM 3.DAT.SG.F help.PST.SG.M
'She wants me to help her.'

Here *-by* is historically part of the irrealis *priechal by* 'would come', the irrealis marker *by* being originally a clitic and therefore occurring in Wackernagel position and fusing with the complementizer. Synchronically, *-by* has become dissociated from the irrealis form of the verb and can also occur, e.g., in infinitival clauses (*čtoby uslyšat'* 'in order to hear'). The resulting irrealis complementizer *čtoby* is used mainly for state-of-affairs complements (as in 80), but it also occurs in the propositional domain: when a propositional attitude verb like 'think' occurring in the matrix clause is negated, the irreality of the proposition that is being rejected may be optionally marked in the form of the complement clause. Examples (81) and (82) show the use of *čtoby* in state-of-affairs and propositional complements respectively:

Russian

(71) *Ja* *ne* *choču,* [***čtoby*** *on* *sjuda* *prichodil*].
1SG.NOM NEG want.PRS.1SG **COMP** 3. NOM.SG.M hither come.IRR.SG.M
'I don't want him to come here.'

Russian

(72) *Ja* *ne* *dumaju,* [***čtoby*** *on* *èt-ogo* *chotel*].
1SG.NOM NEG think.PRS.1SG **COMP** 3.NOM.SG.M this-GEN.SG.N want.IRR.SG.M
'I do not think he wants this.'

Whereas in (81) the irrealis complementizer would also be used without negation, the affirmative variant of (82) could contain only a realis complementizer:

Russian

(73) *Ja* *dumaju,* [***čto*** *on* *èt-ogo* *chočet*].
1SG.NOM think.PRS.1SG **COMP** 3.NOM.SG.M this-GEN.SG.N want.IRR.SG.M
'I think he wants this.'

Latvian, on the other hand, does not collapse these two types of marking. In the counterpart of (71) it has *lai*, the complementizer of state-of-affairs complements, but in that of (82) it will have *ka*, the propositional complementizer, in combination with the irrealis mood (whose function is the same as that of the irrealis complementizer in (82), namely to emphasize the irreality of the proposition that is being rejected):

Latvian

(74) *Es* *ne-gribu,* [***lai*** *viņ-š* *šeit* *atnāk*].
1SG.NOM NEG-want.PRS.1SG **COMP** 3-NOM.SG.M here come.PRS.3
'I don't want him to come here.'

Latvian

(75) *Es* *ne-domāju,* [***ka*** *viņ-š* *to* *gribētu*].
1SG.NOM NEG-think.PRS.1SG **COMP** 3-NOM.SG.M that want.IRR
'I don't think he wants this.'

The coincidence of the oppositions fact vs. state-of-affairs and realis vs. irrealis is therefore merely a default: oppositions in realis vs. irrealis marking may occur both in the propositional (truth-valued) domain and in the state-of-affairs domain. It seems preferable, therefore, to use the terms 'realis' and 'irrealis complementizers' in a way more intimately connected with the use of these terms in the domain of mood. This connection is natural in Slavonic, which has complementizers containing formal irrealis markers (also used as irrealis markers on the verb), but we can also speak of irrealis complementizers with reference to units like Latvian *kaut*, which is semantically (though not formally) marked for irrealis and consistently co-occurs with irrealis marking on the verb. In principle, irrealis complementizers could also be singled out on purely semantic grounds in languages without irrealis marking on the verb.

Note that Lithuanian could be said to have a realis vs. irrealis strategy in complementation, though it is not reflected in the form or choice of the complementizer. There is, however, as noted above (Section 2.1), an opposition between realis and irrealis verb forms which helps identify complement types in the absence of specialized complementizers. Latvian, on the other hand, has no regular correlation between mood and complementation type. Whereas in the case of *kaut*

there is a harmonic combination between the irrealis complementizer and irrealis marking on the verb, with the basic state-of-affairs complementizer *lai* the verb may be realis or irrealis according the 'expectations of actuation'.

The epistemic meanings of complementizers discussed in this article seem to define asymmetric distinctions. There is no epistemic marking correlating with a high degree of confidence in the content of the clausal complement; there is only marking of doubt, or of decisive disagreement. This might be a universal feature: what has been described as a "high degree of credibility" (Noonan 2007: 58) is perhaps only the lack of epistemic qualification. That epistemic marking in complement clauses should be, in a natural way, asymmetric stands to reason. Propositional complements contain propositions that, in simple sentences, are taken to reflect the speaker's propositional attitude. When they are embedded under predicates reflecting other people's utterances or propositional attitudes it is obviously communicatively more important to signal divergence from the speaker's propositional attitude which the proposition would express when used in isolation.

Between the types of distinctions marked in the choice of the complementizers there is, as already mentioned, some degree of overlap. It has been argued in this article that it makes sense to set irrealis complementizers apart from state-of-affairs complementizers for some languages, but for other languages the notions may coincide. Between the notions of proposition vs. state-of-affairs and the modal notions of realis and irrealis there is an obvious connection, although its exact nature deserves to be investigated in more detail.

Between irrealis marking and epistemic qualification in complementation there is a similarity in that both types of marking allow the speaker to qualify the 'reality' of the content of the clausal complement; but only in the case of epistemic marking may this reality be formulated in truth-conditional terms. Irrealis complementizers may only lend an additional shade of irreality to complements that are already inherently non-factive.

9 In conclusion

Though closely related, Lithuanian and Latvian show, from the very start, considerable differences in their use of complementizers. Latvian shows, on the whole, a greater semantic differentiation than Lithuanian: it consistently distinguishes propositional and state-of-affairs complementizers, and among the latter there is even a specialized irrealis complementizer. This semantic marking is partly distinctive, partly harmonic (and thus redundant). Lithuanian makes do with

a smaller number of complementizers, using mood distinctions as a means of opposing clausal complement types, whereas Latvian uses mood to a much lesser extent in contrasting complement types, which leaves realis and irrealis forms of the verb available for expressing 'expectations of actuation'. The main interest of the Baltic complementizer system seems to lie in the insights it affords into the potentially distinct though often by default coinciding oppositions between propositional and state-of-affairs, realis and irrealis, epistemically neutral and epistemically marked complementizers. Also of interest are the recent developments in the domain of epistemic, evidential and interpretive use marking. Markers of these three types acquire specific functions in clausal complements (often differing from the functions they have in independent sentences) and after a period of co-occurrence with complementizers seem to shift to the position of complementizer themselves, leading to the rise of new types of semantically marked complementizers. These processes are still ongoing, especially in Lithuanian, and they are still awaiting more in-depth research.

Acknowledgment

My thanks are due to the editors of this volume as well as to an anonymous reviewer for many useful comments. For any shortcomings of the article I remain solely responsible.

References

Amman, Andreas & Johan van der Auwera. 2004. Complementizer-headed main clauses for volitional moods in the languages of South-Eastern Europe: a Balkanism? In Olga Mišeska Tomić (ed.), *Balkan Syntax and Semantics*, 293–314. Amsterdam & Philadelphia: John Benjamins.

Blakemore, Diane. 2002. *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.

Boye, Kasper. 2010. Reference and clausal perception-verb complements. *Linguistics* 48(2). 391–430.

Bresnan, Joan W. 1979. *Theory of Complementation in English Syntax.* New York & London: Garland Publishing Company.

Dixon, Robert M. W. 2006. Complement clauses and complementation strategies in typological perspective. In Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation*, 1–48. Oxford–New York: Oxford University Press.

Endzelin, Jan. 1923. *Lettische Grammatik*. Heidelberg: Carl Winter.

Frajzyngier, Zygmunt. 1995. A functional theory of complementizers. In Joan Bybee & Suzan Fleischman (eds.), *Modality in Grammar and Discourse* (Typological Studies in Language, 32), 473–502. Amsterdam & Philadelphia: John Benjamins

Fortescue, Michael. 2010. Similitude: A conceptual category. *Acta Linguistica Hafniensia* 42(2). 117–142.

Goodwin, William W. 1900. *Greek Grammar*. Boston: Ginn.

Heine, Bernd & Tanya Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Holvoet, Axel. 2007. *Mood and Modality in Baltic*. Cracow: Wydawnictwo Uniwersytetu Jagiellońskiego.

Holvoet, Axel & Jelena Konickaja. 2011. Interpretive deontics: A definition and a semantic map based mainly on Slavonic and Baltic data. *Acta Linguistica Hafniensia* 43(1). 1–20.

Kavaliūnaitė, Gina. 2008. *Samuelio Boguslavo Chylinskio Biblija. Senasis Testamentas = Vetus Testamentum Lithvanicâ Lingvâ donatum a Samuelo Boguslao Chylinski*, ed. Gina Kavaliūnaitė. Vilnius: Institute for the Lithuanian Language.

Kiparsky, Paul & Carol Kiparsky. 1971. Fact. In Danny D. Steinberg & Leon A. Jakobovits (eds.), *Semantics. An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*, 345–369. Cambridge: Cambridge University Press.

Lichtenberk, František. 1995. Apprehensional epistemics. In Joan Bybee & Suzanne Fleischman (eds.), *Modality in Grammar and Discourse* (Typological Studies in Language 32), 293–327. Amsterdam & Philadelphia: John Benjamins.

Lyons, John. 1977. *Semantics*, vol. 1–2. Cambridge: Cambridge University Press.

Mithun, Nancy. 1995. On the relativity of irreality. In Joan Bybee & Suzan Fleischman (eds.), *Modality in Grammar and Discourse* (Typological Studies in Language 32), 367–388. Amsterdam & Philadelphia: John Benjamins.

Nau, Nicole. 2012. *A Short Grammar of Latgalian*. München: Lincom Europa.

Noonan, Michael. 2007 [1985]. Complementation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol 2. *Complex Constructions*, 2nd edn., 52–150. Cambridge: Cambridge University Press.

Petit, Daniel. 2012. On the Latvian indefinite pronoun *kaut kas*. *Baltic Linguistics* 3. 101–150.

Ransom, Evelyn N. 1986. *Complementation: Its Meanings and Forms* (Typological Studies in Language 10). Amsterdam & Philadelphia: John Benjamins.

Sawicki, Lea. 2012. Responsive discourse particles in Lithuanian dialog. *Baltic Linguistics* 3. 151–175.

Smoczyński, Wojciech. 2007. *Słownik etymologiczny języka litewskiego* [Lithuanian Etymological Dictionary]. Vilnius: Vilnius University.

Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and Cognition*. Cambridge, MA: Harvard University Press.

Stang, Christian S. 1970. Die litauische Konjunktion *jeib* und der lit.-lett. Optativ. In Christian S. Stang, *Opuscula Lithuanica. Ausgewählte Aufsätze und Abhandlungen*, 153–159. Oslo: Universitetsforlaget.

Brian D. Joseph

The semantics and syntax of complementation markers as an areal phenomenon in the Balkans, with special attention to Albanian

1 Introduction

It is well known that various languages of southeastern Europe show a number of grammatical features that have led to them being characterized as a "Sprachbund",[1] i.e., a group of geographically related languages which due to centuries of intense and sustained contact show convergence along structural – as well as, in the typical case, lexical – dimensions. The Balkan Sprachbund group includes, but is not restricted to, languages from five branches of the Indo-European family: Albanian, Greek, the Indic language Romani, the Balkan Slavic languages Bulgarian and Macedonian, and the Balkan Romance (Italic) languages Aromanian, Daco-Romanian, and Megleno-Romanian; the non-Indo-European language Turkish also figures prominently in the Sprachbund, though with less of an effect on structure than on lexis.[2]

Balkan convergent structural features range over all components of grammar, from phonology through pragmatics, but especially noteworthy in the Balkans are features at the level of morphosyntax. There is one morphosyntactic parallel shared by all the languages that is particularly striking, and, in the European context, highly unusual as well. In every one of these languages most if not all complementation – by which is meant clausal structures that fill argument roles in a sentence[3] – involves finite clauses introduced by a subordinating element of some sort and, significantly, containing verbal forms marked for person and number in all of the languages, for aspect in some of the languages, and for tense as well in some. This feature is realized to a considerable extent in each language,

1 The German term *Sprachbund* is commonly used in English, as here, although phrases such as "linguistic area" or "convergence area" are also in use.

2 For overviews of the Balkan Sprachbund, see Friedman (2006), Joseph (2003, 2010); a more comprehensive presentation is to be found in Friedman & Joseph (2017).

3 This definition essentially follows a now-standard view of what complementation is, that given by Noonan (2007: 52): "the syntactic situation that arises when a notional sentence or predication is an argument of a predicate", where serving as subject or object of the predicate is the most usual circumstance.

with Aromanian, Greek, Macedonian, and Romani showing it to the greatest degree, lacking nonfinite complementation altogether.[4] Bulgarian comes close to showing no nonfinite subordination, with just some optional and generally rare use of such nonfinite forms in the standard language, while nonfinite complements play a more significant role in Daco-Romanian and in both major varieties of Albanian, i.e. both the northern Geg and the southern Tosk dialect zones.[5]

Two different types of subordinating element accompany the finite complement in each of the languages: one associated with complements expressing propositions that can have a truth value, here called "indicative" complements,[6] and one associated with nonveridical modality, here called simply "modal". As discussed in Section 2, the indicative complements are introduced by elements – complementation markers – that can be analyzed as canonical complementizers, while the modal complements are accompanied by elements, also a type of complementation marker in that they are associated with complement clauses, that for some of the languages at least are probably best analyzed not as canonical complementizers per se but as mood markers. These elements can be labeled, at least informally, as "dental modal subordinators", abbreviated throughout as 'DMS'.[7] Arguments concerning the status of the DMS markers within the overall class of complementation markers, especially in respect to canonical complementizers, are also presented in Section 2.

Leaving aside Albanian, which is treated in some detail in later sections, some examples of the finite complementation from various languages are given in (1), illustrating both modal and indicative possibilities, presented by examples (i) and (ii) respectively; examples from languages that show nonfinite complementation are given in (2).[8]

---

4 See Section 2 regarding one construction in these languages that under certain definitions not adopted here could be interpreted to show nonfinite complementation.

5 The occurrence of nonfinite complement forms in Tosk and in Geg represents different diachronic trajectories for each dialect: Geg most likely has had infinitives for centuries, possibly since proto-Albanian, and has not undergone as much infinitival replacement as has Tosk (Altimari 2011); Tosk, especially as represented in the Tosk-based now-standard language (in Albanian: *gjuha standarde* 'standard language' or *gjuha letrare* 'literary language') has undergone an expansion in the use of a relatively new infinitive, as discussed below.

6 The characterization "veridical" could also be used for such complements.

7 The term DMS was coined by Victor Friedman, drawing on the adventitious fact that the subordinating element with modal value in all of the languages in question happens to begin with a dental consonant (Albanian *të*, Greek *na*, Balkan Slavic *da*, Balkan Romance *să*, Romani *te*). I say "informally" here so as not to imply that "DMS" is a syntactic category that needs to be recognized in syntactic theory.

8 Here and throughout, the Greek sentences are constructed but are based on my 40 years of work on the language and on examples in the literature or heard in common use. The Albanian

(1) a. i. *Iska-m* [*naši=te* ***da*** *spečel-jat*]. (Bulgarian)
want-1SG ours=ART **DMS** win.PFV-3PL.PRS
'I want our (team) to win.' (literally: "I want that ours win")

ii. *Vjarva-m* [***če*** *naši=te* *šte* *spečel-jat*].
believe-1SG.PRS **COMP** ours=ART FUT win.PFV-3PL.PRS
'We believe that our (team) will win.'

b. i. *emis* *θelu-me* [***na*** *sizitisu-me* *tin* *apofasi* *sas*] (Greek)
we.NOM want-1PL **DMS** discuss-1PL the.ACC decision.ACC your
'We want to discuss your decision.' (literally: "want that we-discuss")

ii. *emis* *pistevu-me* [***oti*** *θa* *nikis-i* *i* *omaða* *mas*]
we.NOM believe-1PL **COMP** FUT win-3SG the.NOM team.NOM our
'We believe that our team will win.'

c. i. *nie* *planira-me* [***da*** *ostavi-me*] (Macedonian)
we.NOM plan-1PL **DMS** leave-1PL
'We are planning to leave' (literally: "planning that we leave")

ii. *misla-m* [***deka*** *Petar* *e* *pameten*]
think-1SG **COMP** Petar is smart
'I think that Petar is smart.'

d. i. *Stă* [***să*** *plou-ă*] (Daco-Romanian)
is.about.3SG **DMS** rain-3SG.SBJV
'It is about to rain.' (literally: "about that it rain")

ii. *Mi=a* *spus* [***că*** *e* *supărat*]
me.DAT=has.3SG told **COMP** is.3SG angry
'He told me that he was angry.'

(2) a. *toj* *ne* *smea* [*se* ***obadi***] (Bulgarian)
he.NOM NEG dares.3SG REFL **answer.INF**
'He does not dare answer'.

b. *pare* [***a*** ***ploua***] (Daco-Romanian)
seems.3SG **INFM** **rain.INF**
'It seems to be raining'.

The inclination toward exclusively finite complementation structures actually represents the synchronic result of two different but related diachronic developments: a reduced use of forms populating a previously existing category of infini-

examples are mostly taken directly or adapted from standard sources, especially Newmark, Hubbard & Prifti (1982) but also Camaj (1984), though some were found on-line via a targeted google search. The Bulgarian examples have been checked with a native speaker; I thank Dr. Anastasia Smirnova of the University of Michigan for her help here; example (2a) is from a standard grammar of essentially 20th century literary usage and sounds unnatural or dialectal to most speakers today. Macedonian examples are adapted from standard sources. The Daco-Romanian and Aromanian examples come from standard reference works, especially (Pană Dindelegan (2013) for Daco-Romanian and Vrabie (2000) for Aromanian. Note that the Romanian infinitive is introduced, in most contexts in which it is used, by an element *a*, which for want of a better classification and in the absence of careful analysis is simply labeled "INFM" here, for "infinitive marker" (also used here with Albanian elements).

tive and an expanded use of finite verbs in subordination. The diachrony, while interesting in its own right, is not relevant here, as the focus here is on the synchronic nature of complementation in the Balkans but more particularly on the syntax and especially the semantics of the complementizer and complementizer-like elements that occur in complement structures. Although the language focus is on the Balkans in general, particular emphasis is placed on Albanian, as it is perhaps the least well described of the languages of Europe that serve as a national (official) language of a nation-state,[9] as far as modern analytic accounts are concerned. I turn first, therefore, to some general facts about the syntax of complementation in the Balkans and then home in on Albanian, both as to syntax and semantics, with attention to other Balkan languages where appropriate.

2 On the syntax and morphosyntax of Balkan complementation

The occurrence of finite complementation in these languages has consequences for the syntax of complement clauses. Moreover, there are two issues of a more morphosyntactic nature that need to be addressed. Thus before moving into a consideration of the semantics of Balkan complementation and into Albanian per se, some discussion of the syntax and morphosyntax is in order.

First, as to the verb forms themselves that occur in the complement clauses, for the indicative clauses there is nothing particularly noteworthy: the full range of verb forms that can occur in main clauses is available for use in Balkan indicative complement clauses – present tense, future tense, perfect forms, and any of the various past tenses for those languages that, for instance, distinguish imperfect (past imperfective) from aorist (past perfective). All of these can occur in subordinate clauses with the indicative complementizer (e.g. Bulgarian *če*, Greek *oti*, Macedonian *deka* etc.), embedded under suitable matrix verbs; in place of the subordinate verb in the indicative sentences in (1), for instance, other indicative verbs could be used, as in the Greek examples in (3) – similar examples can be constructed for the other languages.

(3) a. *emis* *pistevu-me* [***oti*** *nikis-e* *i* *omaða* *mas*]
we.NOM believe-1PL **that** won.PFV-3SG the team.NOM our
'We believe that our team won.'

9 Albanian is the official language of both Albania and Kosovo.

b. *emis* *pistevu-me* [***oti*** *nikus-e* *i* *omaða* *mas*]
we.NOM believe-1PL **that** won.IPFV-3SG the team.NOM our
'We believe that our team was winning.'

c. *emis* *pistev-ume* [***oti*** *ex-i* *nikisi* *i* *omaða* *mas*]
we.NOM believe-1PL **that** has-3SG win.PRF the team.NOM our
'We believe that our team has won.'

d. *emis* *pistevu-me* [***oti*** *θa* *nikuse* *i* *omaða* *mas*], (*an* ...)
we.NOM believe-1PL **that** FUT win.3SG.IPFV the team.NOM our if
'We believe that our team would win (if ...)'

The indicative complementizer seen here in Greek, *oti*, is used for nonfactive complements; it varies with another form *pos*, most likely derived from the question word for 'how' (also *pos*), which covers the same semantic range as *oti* so that the choice between the two seems to be a matter of stylistics. For factives, where the truth of the complement is presupposed, Greek uses a different indicative complementizer, *pu*, perhaps derived from the earlier locative relative *(h)opou* 'where, (in a place) in which'. Thus the Greek examples in (3) can have *pos* instead of *oti* but not *pu* (so that *emis pistevume pos nikise i omaða mas* is an acceptable variation on (3a) but **emis pistevume pu nikise i omaða mas* is not). An acceptable example of complementation with *pu* is given in (4), with the factive verb *ksexno* 'forget'.

(4) *ksexas-a* [***pu*** *i-ne* *jatro-s* *o* *petro-s*]
forgot-1SG **COMP** is-3SG doctor-NOM the.NOM Peter-NOM
'I forgot that Peter is a doctor.'

In (4), the truth of the complement is not in question, i.e. Peter is indeed a doctor. Other factive complement-taking-verbs in Greek include *lipume* 'be sorry', *xerome* 'be glad', and *ime perifanos* 'be proud', inter alia. Other Balkan languages seem not to have a specialized complementizer for factivity; in Albanian, for instance, the indicative complementizers *që* and *se* are used, essentially interchangeably, with both factive and nonfactive complements:

(5) a. *Më* *vjen* *keq* [***që*** *vdiq-ën* *dy* *njerëz* *të rinj*]
me.DAT come-3SG bad **COMP** died-3PL two men young
'I am sorry that two young men died.' (literally: "It-comes to-me badly that ...")

b. *Më* *vjen* *keq* [***se*** *përfundov-e*]
me.DAT come-3SG bad **COMP** finished-2SG
'I am sorry that you finished.' (literally: "It-comes to-me badly that ...")

In the modal complement clauses, however, most of the languages show a restriction of some sort regarding which verb forms can occur. In particular, in all of them, modal clauses allow for the occurrence of special verb forms that cannot occur, or do not freely occur, as such independently in main clauses. This restric-

tion takes a different form in the different languages, and, in certain ways, is not unlike the occurrence in some languages of western Europe (and elsewhere) of special "subjunctive" verb forms that are tied to subordination. In fact, in Albanian and in the Balkan Romance languages, there are some morphologically distinct mood forms[10] to be found that are traditionally referred to as "subjunctive",[11] and, with certain definable exceptions,[12] they occur only in subordinate clauses with a DMS; some examples showing these forms specifically in Albanian, which are morphologically distinct from indicative forms generally only in 2nd and 3rd person singular forms in the present subjunctive,[13] are given in (6):

(6) a. *Dua* [***të*** ***shko-sh*** *në* *Shqipëri* *dikur*].
want.1SG **DMS** **go-2SG.SBJV** to Albania sometime
'I want you to go to Albania sometime.'
(literally: "I-want that you-go ...", cf. *shko-n* 'go-2SG.IND')

b. *Dua* [***të*** ***ke-të*** *Agim=i* *dhurat-ën*].
want.1SG **DMS** **have-3SG.SBJV** Agim=ART.NOM gift-DEF.ACC
'I want Agim to have the gift.'
(literally: "I-want that he-have ..."; cf. *ka* 'has.3SG.IND')

The examples in (7) show that these particular forms, *tout court*, are excluded in main clauses:

(7) a. **Shko-sh* *në Shqipëri* / ok*Shko-n* *në Shqipëri*
go-2SG.SBJV go-2SG.IND

10 See footnote 11 on the Albanian forms; in Daco-Romanian there are distinct subjunctive endings only in the 3rd person singular and plural, whereas Aromanian has special subjunctive forms only for *hiu* 'be', *am* 'have', and *ştiu* 'know'.

11 The native term in Albanian is *lidhore* (from *lidh* 'tie, connect'), and in Daco-Romanian *conjunctiv*. Albanian actually has a mood system that, from a morphological standpoint, is quite well developed, with more distinct categories than any other Balkan language. Besides an indicative and subjunctive, Albanian has a morphological optative, e.g. *rrofsh* 'may he live' (root *rro-* 'live'), an imperative, e.g. *jini* '(you all) be!' (root *je-*), and admirative (marking nonconfirmativity), consisting of a truncated form of the perfect participle fused with an inflected form of 'have' e.g. *qenke* '(I can't believe) you are ...! / are you really...?!' (root *qe-* 'be', and cf. *ke* 'you have').

12 For instance, in Albanian colloquial usage, the DMS *të* can be omitted in future forms, so that, for instance, *do shkosh* is an acceptable variant of *do të shkosh* for 'you will go'. It is admittedly debatable as to whether even the fuller form, *do të shkosh*, involves subordination; that is, it would take a fairly abstract analysis, with *do*, as the marker of futurity, being represented as a higher verb that takes *të shkosh* as its complement. Such analyses have been proposed, and were especially popular in the era of Generative Semantics, but this is not the place to engage in this debate further.

13 The auxiliary verbs *jam* 'be' and *kam* 'have' have distinct subjunctive forms also in the 1SG and 3PL forms (*jem*/*kem* and *jenë*/*kenë*, respectively). Albanian can be said to have a past and a perfect subjunctive in the analysis of *të* given below.

b. **Ke-të Agimi dhuratën* / *[ok]Ka Agimi dhuratën*
have-3SG.SBJV have.3SG.IND

In addition, other forms that are identical with the indicative, especially the imperfective past ("imperfect") but also, with the subjunctive forms of 'have', a perfect formation (and pluperfect and future perfect, through the availability of a past and future of 'have', respectively) can occur in DMS clauses together with *të*.

In Greek and Balkan Slavic (Bulgarian and Macedonian), however, the restriction is tied to aspect: present perfective aspect forms can never occur as bare forms in main clauses, i.e. without some sort of "supporting" element; they are always accompanied by, and actually introduced by, a marker of some sort, most notably the future tense marker or, in Greek, a hortative marker, as in (8).

(8) a. **akus-o kaθara tora ja proti fora*
hear.PFV-1SG.PRS cleanly now for first time
Intended: 'Now, for the first time, I hear clearly.'
b. *θa akus-o prosextika apo tora*
FUT hear.PFV.1SG.PRS carefully from now
'I will listen carefully from now on.'
c. *as akus-o prosextika tora*
HORT hear.PFV-1SG.PRS carefully now
'Let me listen carefully now.'

Imperfective forms are not subject to such a restriction, as a comparison of (8a) with (9) shows:

(9) *aku-o kaθara tora ja proti fora*
hear.IPFV-1SG.PRS cleanly now for first time
'I hear clearly now for the first time'

In subordinate clauses of various sorts, as in (10a), such perfective forms will always have some sort of "supporting" element, e.g. a subordinating conjunction or an indefinite relative word, and in modal complementation, as in (10b), the DMS provides the support:

(10) a. *an akus-o* / *otan akuso* / *opjon akuso*
if hear.PFV-1SG.PRS when hear.PFV-1SG.PRS whomever hear.PFV-1SG.PRS
'if I hear …' / 'when I hear …' / 'whomever I hear …'
b. *ipósxo-me* [***na*** *akus-o prosextika*]
promise-1SG **DMS** hear.PFV-1SG.PRS carefully
'I promise to listen carefully.'

There is more that can be said about the distribution of these verb forms. However, since the attention in this study is primarily on the semantics of complementizers and complementation, with Albanian as the main language focus, further discus-

sion of these restrictions can be left to other venues,[14] and other morphosyntactic and syntactic issues that Balkan finite complementation raises can be addressed.

A second syntactic issue has to do with the characterization of the element (for want of a better neutral term) that introduces the complementation. As suggested in Section 1, the issue is whether all the elements that introduce or are associated with complementation in these languages, that is forms like *deka*, *oti*, *da*, *na*, etc., are complementizers or are instead something else. This question assumes that there is some empirical content to the designation "complementizer", some consequences that follow from this morphosyntactic category label. "Complementizer", following Noonan (2007: 55) can be taken to be "a word, particle, clitic, or affix, one of whose functions is to identify the entity [i.e., a complement type] as a complement", i.e. a notional sentence that fills an argument role (see footnote 3). This is admittedly a broad definition, but it focuses, quite properly, on those elements that allow a clause to function as a complement. However, for head-initial (right-branching) languages like those in the Balkans, such elements can be further differentiated by an added characteristic that at least some complementizers in other languages show, namely that of sharply delimiting a clause boundary.[15] Admittedly, this is not a part of Noonan's now-standard definition, but it is consistent with the function served by elements given the label "complementizer" in various accounts of diverse languages. This added characterization of complementizer introduces some gradience into the identification of complementizers,[16] but with it, the indicative subordinators like the Greek *oti* and Macedonian *deka* can be recognized as true, i.e. canonical, complementizers as they do not allow pieces of the clause they introduce to "leak" and occur

14 See Joseph 2012 for some relevant discussion and references regarding Greek.

15 Compare English *that*, which allows topicalization by fronting of a complement-clause constituent within its bounds but not to its immediate left, as in the acceptable (i) compared with the unacceptable (ii):

(i) *John stated that for nourishment nothing beats beans*.
(ii) **John stated for nourishment that nothing beats beans*.

I say "immediate left" because typically, subordinate clause elements can be fronted to the left periphery of the main clause, in, e.g., a topicalized or focalized construction.

16 Note that Noonan (2007: 55), despite recognizing affixes as possible complementizers (see above), excludes "derivational affixes, like English *-ing*, which are used to convert a form from one part of speech to another". The suffix *-ing*, however, allows a verb, through the creation of a deverbal noun, to serve in an argument role; cf. *Walking daily is healthy*, where *walking* is comparable to an infinitive (*To walk daily is healthy*), which Noonan does see as a complement (and *to* thus as a complementizer). Noonan's decision ultimately to treat *-ing* forms under the rubric of complementation (Noonan 2007: 118) may reflect a realization on his part of complementizer-like properties, under his functional definition, of these nominalized forms.

to their immediate left,[17] as shown in (11), for instance, to illustrate this property just with Greek:

(11) **emis* *pistevu-me* *i* *omaða* *mas* ***oti*** *nikis-e*
we.NOM believe-1PL the.NOM team.NOM our **that** won-3SG.PST
Intended: 'We believe that our team won.'

By this test, the DMS elements found in the various Balkan languages are not as fully complementizer-like as the indicative subordinators. Again illustrating from Greek, sentences such as (12) show that the DMS *na* does not demarcate a clause boundary, since subordinate-clause elements can be positioned on its immediate left:[18]

(12) *emis* *elpizu-me* *i* *omaða* *mas* ***na*** *nikis-i*
we.NOM hope-1PL the.NOM team.NOM our **DMS** win.PFV-3SG
'We hope that our team will win'.

This issue has been much discussed in the literature on Greek. Most recently, Sampanis (2011, 2013) argues that *na* is a mood marker and not a complementizer, stating (2013: 168): "the *na*+verbal form configuration is a genuine mood. The particle *na* is an affix-like inflectional element and the host of the semantic features *-assertion*/*-realis*. In consequence, any analysis in terms of derivational syntax should take into account the morphological and the semantic properties of the M[odern]G[reek] subjunctive; hence the "particle" *na* should not be considered a complementiser." Parallel considerations hold for the DMS elements in the other languages, with similar kinds of evidence available to be brought to bear on the question.[19]

This analytic decision concerning the status of the DMS means that DMS clauses, while functionally being able to serve as complementation in that they fill

17 Regarding the need to specify "immediate left", see footnote 15.

18 Noonan (2007: 77, 80) labels Albanian *të* with "COMP" (i.e., a complementizer) due to his reliance on a functionally based definition; his example (114) has ordering parallel to that in (12), with a subordinate clause subject to the immediate left of *të*.

19 There are other interpretations possible here. It has been suggested by Petar Kehayov, for instance, that the DMS could indeed be a complementizer but one with narrower scope, introducing not propositional complements, but rather only complements that designate events (states-of-affairs). As such, one could say that it must occur closer to the core of a sentence's predication, that is, to the verb. I welcome such added functional perspectives but for here, choose to stay with the use of a formal criterion to characterize a key way in which the indicative subordinators in the Balkans and the DMS elements differ.

the object argument slot for verbs like 'want', 'try', 'hope', 'begin' and so on,[20] do not have a complementizer per se introducing them; this fact can be represented formally by positing a null ("zero") complementizer that heads the DMS clauses in the usual case. This last qualifier is needed because in some instances, DMS clauses can co-occur with canonical complementizers; in Greek, for example, the factive complementizer *pu* 'that', as seen in (4), repeated here as (13a), can also introduce restrictive relative clauses (much as *that* does in English), as seen in (13b), and can be followed by a *na*-clause, as seen in (13c).

(13) a. *ksexas-a* [***pu*** *i-ne* *jatro-s* *o* *petro-s*]
forgot-1SG **COMP** is-3SG doctor-NOM the.NOM Peter-NOM
'I forgot that Peter is a doctor' (NB: this entails that Peter is a doctor)

b. *o* *tipo-s* ***pu*** *bik-e* *i-ne* *o* *filo-s* *mu*
the.NOM guy-NOM **COMP** entered-3SG is-3SG the.NOM friend-NOM my
'The guy who came in is my friend'

c. *psaxn-o* *kapjo-n* ***pu*** ***na*** *me* *voiθis-i*
search-1SG someone-ACC **COMP** **DMS** me.ACC help-3SG
'I am looking for someone that might help me'.

In (14a), an example from Albanian that matches (13c) is given, where *që* is parallel to Greek *pu* as an invariant form that is able to introduce relative clauses, though (14b) offers a case where a clause that is the complement to a nominal predicate is introduced by *që* together with a co-occurring DMS:[21]

(14) a. *Po* *kërko-ja* *ndonjë* ***që*** ***të*** *ish-te* *pak* *i jashtëm*
PROG seek-1SG.IPFV someone **COMP** **DMS** was-3SG little outside
'I was searching for someone who might be a bit of an outsider'

b. *Por* *detyra* *ime* *si* *komunist* *ësh-të* ***që*** ***ta***[22]
but duty.NOM my as Communist be-3SG **COMP** **DMS;it.ACC**
shpejto-j *këtë* *ditë*
accelerate-1SG this day
'But my duty as a Communist is to accelerate this day.'
(literally: "...duty ... is that I might accelerate ...")

Here we have another reason to treat the DMS markers as being less complementizer-like than the indicative subordinators: DMS can themselves co-occur with

20 DMS clauses can fill other functions, such as the expression of purpose, hence the phrasing "being able to serve".

21 The combination *që të* also occurs in the meaning of 'so that', i.e., in an adjunct, not argument, usage, though (14b) seems not to show that meaning.

22 The form *ta* is a portmanteau combination of DMS *të* and the weak object pronoun *e*.

complementizers (such as Greek *pu* or Albanian *që*);[23] see the end of Section 3 for more on combinations of these elements.

The analysis of the DMS as a mood marker combines with the range of verb forms that can appear with the DMS elements to give a rather elaborated set of moods in the various languages. As noted in footnote 13, Albanian is thus said to have not only a present tense subjunctive but also a past subjunctive and a perfect subjunctive (so Newmark, Hubbard & Prifti 1982; Camaj 1984); the same could be said about Greek, as noted in Joseph (2012), although few linguists and grammarians have taken that analytic step.

Moving beyond morphosyntax and node labels, there are other aspects of the syntax of complementation that are affected by the pervasive occurrence of finite complementation in each of the Balkan languages. Most importantly, while there is no overt nominative subject in the complement clauses in the example sentences in (1), these verbs can in principle, under appropriate conditions of emphasis or contrast, occur with an overt subject nominal that is in the same form as in a main clause, e.g. nominative case in the languages that distinguish case in the nominal system; thus, (15a) in Greek is an acceptable variant of (1bi), repeated here as (15b) for convenience of reference.[24]

(15) a. *θelu-me* [***na*** *sizitisu-me* ***'mis*** *tin* *apofasi* *sas*]
want-1PL **DMS** discuss-1PL **we.NOM** the.ACC decision.ACC your
'We want to discuss your decision'

b. *emis* *θelu-me* [***na*** *sizitisu-me* *tin* *apofasi* *sas*]
we.NOM want-1PL **DMS** discuss-1PL the.ACC decision.ACC your

This means that when a complement clause in Greek or another Balkan language with finite complementation lacks an overt subject, this absence is due to whatever process or processes allow for the absence of unemphatic subjects in surface clauses, and not some special process associated with complementation.[25]

23 There are of course languages, such as earlier stages of English that allow for "doubly-filled" COMP nodes, but beyond this combination of *pu* with *na*, there does not seem to be any independent justification for such a construct in Greek. In particular, *pu oti* is not a possible combination (nor is *që se* in Albanian), so there is this additional point of differentiation between the DMS elements and the indicative subordinators.

24 The form *'mis* in (15b) with the initial *e-* elided is a phonological variant of *emis*, the use of which is dictated by its occurring after the vowel-final ending of *θelume*.

25 Noonan (2007: 75–79) discusses the absence of subjects with complement verbs under the rubric of "Equi-deletion", drawing on classical Transformational Generative Grammar notions and terminology and illustrating the phenomenon with examples such as English *Zeke wants to plant the corn*. However, he is careful to distinguish the Balkan finite-complement situation, using Albanian sentences as his examples, and opts instead for an analysis of it in which the absence of the complement-clause subject "follows the usual discourse conditions on anaphoric

There is a further consequence of this interpretation of complement structures and complementizers. Complementation always involves two verbs in a particular relation to one another, with one verb heading the main clause and one heading the subordinate clause. Moreover, there are combinations of two verbs for which one might propose a complementation relation since one verb "completes" the argument requirements of another verb. For instance, all of the Balkan languages have formations with 'have' or 'be' plus a participial form that are past tenses or perfect tenses, e.g. Albanian *kam ardhur* 'I-have come' or Macedonian *rešil sum* 'I solved (literally, "solved I-am"), and one could view these, at a somewhat abstract level perhaps, as involving some kind of main-verb/subordinate-verb relation, with, e.g., *ardhur* completing the combinatoric needs of *kam* in this particular construction.[26] However, with the view of how to identify complementation and complementizers taken here, since there is no subordinating word that delimits the break between the two verbs, i.e. no overt complementation marker, these are to be analyzed as having no complementizer node at all occurring with the second verb. The second verb would thus be an instance of a Verb(Phrase) rather than a ComplementizerPhrase (or Sentence),[27] so that these combinations would not represent complementation but instead should be taken as auxiliary-plus-main-verb constructions, i.e. with a somewhat "flat" linear structure rather than a more hierarchical one with a level of clausal embedding.

Given that DMS clauses do not involve overt complementizers in the analysis adopted here, extending the scope of this chapter to clauses with complementation markers in general as opposed to just complementizers allows for the semantically very interesting DMS clauses to be included. That is, if one were to focus just on the semantics of complementizer elements per se, then DMS clauses would not get any attention inasmuch as they lack an overt complementizer. It is only once the syntactic argumentation is made, as here in Section 2, that DMS clauses do contain a complementizer position or node but one that is unfilled,

ellipsis and is not the product of a sentence-internal process like equi" (Noonan 2007: 77–78). In a syntactic theory such as the Principles and Parameters model (also known as Government and Binding; Chomsky 1981, 1982), a missing subject pronoun with a finite verb, e.g. *θelume* in (15a), would be described as involving "Subject *pro*-Drop" or as adhering to the "Null Subject Parameter"; moreover, with regard to the different null subjects recognized in that framework – *pro* ("little *pro*") for the "dropped" null subject and PRO ("big PRO") for "Equi-deletion" cases (more usually referred to now rather as "control" cases) – complement clauses like that in (15b) would have *pro* and not PRO.

26 Such analyses were adopted within the framework of Generative Semantics in the late 1960s and early 1970s, the *locus classicus* being Ross 1969.

27 This means distinguishing between the absence of a complementizer node altogether, as here, and the presence of such a node but with nothing filling it, as proposed for DMS clauses.

and are thus with a "zero" or "implicit" complementizer, that the semantics of DMS clauses become relevant for complementizer semantics, through an examination of the semantics of the zero complementizer plus the DMS clause.

Finally, it is important to realize that the notions "complementation" and "complementizer" overlap but are not coextensive. That is, complement clauses fill argument positions, as noted at the outset in Section 1, but complementizers, i.e. delimitors of clause boundaries, introduce not just argument clauses but also adjunct clauses. Clause-demarcating words are found with adjuncts too, and express various sorts of semantic functions, including temporal relations, concession, conditions, and the like. These functions are discussed, where appropriate, in the sections that follow, serving as reflections of the range of complementizer semantics that go beyond complementation. Other ways in which complement clauses go beyond simple complementation are also presented.

3 Complementation in Albanian

It is appropriate at this point to offer more perspective on Albanian as it allows for the exploration within a single language of the two types of complement structure – finite and nonfinite – found in the Balkans. This is due to the fact that Albanian as a language is one of those in the Balkans that shows both finite and nonfinite complementation, both as a key isogloss separating the Geg and Tosk dialects and as variation within both dialects but especially within the generally Tosk-based standard language. Accordingly, examples of infinitival usage from Geg are given in (16) and from Tosk in (17), with finite complement counterparts given in (18) for Geg, where such variants are possible, and in (19) for Tosk (cf. example [6] above for other Tosk examples).

(16) a. *fillo-va* [***me*** ***qeshun***]
begin-1SG.PST **INFM** **laugh**
'I began to laugh.'

b. *ka* *pasë* *fat-in* [***me*** ***njohun*** *shumë* *njerëz* *të* *letrave*]
has.3SG had luck-DEF.ACC **INFM** **know** many men of letters
'he has had the good fortune to know many men of letters'

c. *mund* *të* *je-të* *e vështirë* [***me*** ***ngrënë***] *dhe* [***me*** ***pi***]
can DMS be-3SG.SBJV difficult **INFM** **eat** and **INFM** **drink**
'It can be difficult to eat and to drink.'

d. [***me*** *i* ***ikë*** *rrezik-ut*] *nuk* *âsh-t* *ligështí* *por* *urti*
INFM it.DAT **leave** risk-DEF.DAT NEG be-3SG cowardice but prudence
'To avoid danger is not cowardice but prudence.'
(Camaj 1984: 247)

(17) a. *Mos* *luftërat* *sherbejnë* [***për të*** ***shitur*** *armë*]?!
NEG.Q wars serve.3PL **INFM** **sell** weapons
'Do wars not serve to sell weapons?!'
b. *propozimi* [***për të*** ***ndryshuar*** *pjes-ën* *e rezolutës*][28]
the.proposal **INFM** **change** part-DEF.ACC of.the.resolution
'the proposal to change part of the resolution'
c. *ësh-të* *e vështirë* [***për të*** ***thënë***]
be-3SG.PRS difficult **INFM** **say**
'It is difficult to say.'

(18) *mund-em* [***ta*** *mbaro-nj*]
can-1SG.NACT **DMS;it.ACC** make-1SG
'I can make it.'(Camaj 1984: 247)

(19) a. *po* *fillo-ni* [***të*** *kupto-ni* *tani*]
PROG begin-2PL **DMS** understand-2PL now
'You are beginning to understand now.'
b. *do* *të* *vazhdo-jmë* [***të*** *ul-emi* *atje*]
FUT DMS continue-1PL **DMS** sit-1PL.NACT here
'We will continue to sit here.' (literally: "We will continue that we sit here")
c. *ësh-të* *e vështirë* [***të*** *qesh-in*]
be-3SG.PRS difficult **DMS** laugh-3PL
'It is difficult for them to laugh' (literally: "It is difficult that they laugh")
d. *mund* [***të*** *shko-ni* *në* *Tiranë*]
can **DMS** go-2PL to Tirana
'You(-all) can go to Tirana' (literally: "It-can that you(-all) go to Tirana")

For the most part, these complement clauses fill object argument slots, though the Geg infinitive offers the possibility of a subject complement occurring preverbally in sentence-initial position, as (16d) shows. Besides these DMS complements, indicative complementation with *se* is also to be considered; examples include (20) from Geg and (21) from Tosk (cf. also [5b] above for another example).

(20) a. *plak-a* *i* *tha* [***se*** *kish-te* *bâ* *gadi* *gjithshká*]
old.woman-DEF him.DAT said.3SG **COMP** had-3SG made ready all
'The old woman said that she had prepared everything.'
(Camaj 1984: 247)
b. *u* *bë* *mirë* [***se*** *u* *takua-m* *pasdrekë*]
NACT made.3SG good **COMP** NACT met-1PL afternoon
'It was (literally: "became/was-made") good that we met in the afternoon'.
(Camaj 1984: 246)

28 This example has an infinitival clause dependent on a noun, but since the noun *propozim* is a nominalization from the verb *propozoj* 'propose', which takes a DMS-clause complement, this can be considered to be complementation in the sense adopted herein.

(21) a. *mendo-j* [***se*** *Saranda ësh-të në jug*]
think-1SG.PRS **COMP** Saranda is-3SG in south
'I think that Saranda is in the south'

b. *beso-jnë* [***se*** *këto probleme nuk zgjidh-en lehtë*]
believe-3PL.PRS **COMP** these problems NEG solve-3PL.NACT easy
'They believe that these problems cannot be solved easily'

In (20b), the *se*-complement is positioned post-verbally ("extraposed") but is still arguably filling a subject slot; unlike the subject infinitive of (16d), however, preverbal positioning is not possible here (cf. **Se u takuam pasdrekë u bë mirë*).

The sentences given here in (16) through (21) thus go together with the earlier examples in (1)–(4) from other languages to demonstrate the Balkan realization of indicative and modal finite complementation in Albanian, though with the infinitival possibility as well. Infinitival usage is more prevalent in Geg than in Tosk, and accordingly, DMS clauses are more frequent in complement positions in Tosk than in Geg.

It is noted above, *à propos* of examples (13) and (14), that certain combinations of subordination markers are possible. For instance, *pu na* occurs in Greek, though *pu* here is probably better taken as a relative marker per se and not as a factive complementizer. And, in Albanian, the combination *që të* occurs; recall the examples in (14). The combination *oti na* is not possible in Greek, nor *pos na* (except if it is the question-word *pos* 'how?'), nor *pu oti* (cf. footnote 23). It is interesting to note further that the Albanian form *se* can co-occur with modal *të*, as in (22), in what looks on the face of it like the indicative complementizer joining up with the subjunctive marker.[29]

(22) a. *para* ***se të*** *bi-nte* ...
before ***se*** **DMS** fall-3SG.IPFV
'Before he fell ...' (literally: "before that that he fell")

b. *Përpara* ***se të*** *nise-jë nga Kanadaja drejt Tiranë-s* ...
before ***se*** **DMS** set.out-3SG.SBJV from Canada to Tirana-DAT
'Before he sets out from Canada to Tirana, ...'

Such examples are problematic if *se* is the indicative (canonical) complementizer, for the usual (indicative) mood-selecting properties of *se* would seem to be at odds with the subjunctive modality of *të*. A closer consideration, however, suggests that *se* here has a quite different function, as it is combining with a certain subordinator (the DMS *të*) to introduce adjunct clauses, not the argument clauses introduced by the indicative *se*. In fact, it is especially common to see *se të* in these expressions for 'before', and it may only be a matter of convention that *(për)*

29 No complement-marking brackets are given in (22) owing to the ambiguities in the analysis of *se*.

para se is written as two words, thus inviting an analysis of *se* as the indicative complementizer. Interestingly, the conjunction *megjithëse* 'although', as in (23):

(23) *Megjithëse* *të* *je-mi* *të sinqertë*, ...
although.*se* DMS are-1PL sincere
'Although we are sincere, ...'

is conventionally written (now at least) as one word, even though, from an historical perspective, it is built up of *me* 'with' + *gjithë* 'all' + a form *se*, presumably, but not necessarily, at one time the indicative complementizer. Such combinations may date to a time when *se* had a wider functional range than it does now and may thus show complementizer *se* co-occurring with modal *të* only from an historical perspective and not synchronically for contemporary Albanian.

4 The semantics of complementation in Albanian

The previous sections offer the essential background for viewing the specifically semantic properties of complementation in Albanian, as representative of the situation in the Balkans more generally. As the detail given in those sections shows, there is a considerable amount of morphology and syntax needed as background in order to make sense of complement structures, but with that information in place, the semantic distribution can be discussed more meaningfully.

As has been emphasized already, the key distinction for Albanian (and Balkan) complements is modal versus nonmodal (indicative), as seen in complements with a DMS versus those with canonical complementizers *oti*/*că*/*se*/etc. Thus, the key semantic element lies in this distinction. Indicative complementation, as Newmark, Hubbard and Prifti (1982: 78) put it, "affirms the independent reality of the action". It is probably fair to say that the indicative complement clauses do not have particularly unusual – and therefore, in a sense, not particularly interesting – semantics, inasmuch as they always present information that is rooted in the real world, dealing with the observable and the knowable. Verbs that govern indicative complements include *besoj* 'believe', *mendoj* 'think', *njoftoj* 'inform', *them* 'say', and the like, that is, verbs of assertion and the expression of a propositional attitude; the complements then provide the content of the assertion or the focus of the attitude. In this way, all of these verbs govern complements that have a truth value, so that their semantics are tied in some way to realia; the propositions that are reported with such verbs can of course, however, range over the fantastic and the unreal.

With the modal complements, by contrast, the semantics begin to get interesting, as a wider range of meanings must be considered. Newmark, Hubbard and Prifti (1982: 78) describe the essential use of the Albanian subjunctive, i.e. clauses with the DMS *të*, in the following way: "[t]he subjunctive mood is basically the mood that indicates dependency of the verb. In most sentences a subjunctive verb form will be preceded by an antecedent modal, verb, adjective, adverb, conjunction, noun, or particle to which it is subjoined. [...] Using subjunctive forms [...] express[es] possibility, desirability, or obligation". Given that such DMS clauses are syntactically dependent, their semantics in part depend on the particular combination of controlling predicate plus *të* clause. Some examples of such semantic (and syntactic) dependency are given above in (18) and (19), with controlling verbs that are modal (*mund(em)* 'can') or phasal (*filloj* 'begin', *vazhdoj* 'continue'), but a clearer sense of the semantics of these complement clauses comes from a fuller description of the range of predicates that subjunctives can be dependent upon and that they complete and thus combine with. Predicates that take the subjunctive include modal predicates that convey obligation, with examples in (24a), desiderative predicates that express a speaker's will, as given in (24b), manipulative predicates that express the imposition of will, exemplified in (24c), and those that express a speaker's inner state, as listed in (24d):

(24) a. *duhet* 'must, need, should', *është e nevojshme* '(it) is necessary', *ka nevojë* 'there is need', *lipset* 'must, need, should'
b. *dëshiroj* 'desire', *kërkoj* 'seek', *përpiqem* 'try', *synoj* 'intend', *uroj* 'congratulate, wish well'
c. *këshilloj* 'advise', *lejoj* 'permit', *lut* 'pray', *ndaloj* 'forbid', *porosit* 'request', *urdhëroj* 'order'
d. *di* 'know how', *dua* 'like, love', *dyshoj* 'doubt', *guxoj* 'dare', *mendoj* 'think', *nuk besoj* 'disbelieve', *më pëlqen* 'like' (literally: "to-me it-pleases"), *preferoj* 'prefer', *pres* 'expect', *vendos* 'resolve, decide'

Comparing indicative complementation with modal complementation, then, one can see that the indicative combinations generally involve propositional complements while the modal combinations designate states-of-affairs. From a cross-linguistic perspective, this range of governing predicates and thus this range of semantics for modal and indicative complementation is in itself perhaps not unusual. Still, the reliance, in Tosk especially, on finite means of expressing such complement modality makes Albanian, along with the other Balkan languages, somewhat unusual, particularly in the European context where propositions tend to go with finiteness and states-of-affairs with nonfiniteness (Cristofaro 2003).

The subjunctive can also be used without a governing predicate, also by way of expressing various kinds of modality. This type of subjunctive use, found as well in other Balkan languages, is taken up in Section 5.

One phenomenon that is mostly restricted to Albanian among the Balkan languages that affects the semantics of complement forms is a sequencing of tenses when DMS complements are involved. Thus, in (25), the imperfect subjunctive is the norm in standard/literary usage with a past tense main verb.

(25) *ai* *vendos-i* [***të*** *qëndro-nte*]
he.NOM decided-3SG **DMS** stay-3SG.IPFV
'He resolved to stay.' (literally: "He resolved that he stayed'")

A form of (25) with the present subjunctive (*ai vendosi të qëndrojë*) is possible colloquially (Newmark, Hubbard & Prifti 1982: 80). Nonetheless, the pattern of tense-matching as in (25) is quite regular in Albanian and is apparently the source of a parallel construction found in Aromanian. Sandfeld (1930: 117–118) cites the Aromanian example in (26):

(26) *cu* *vruta* *nu* *putea-m* *s* *me* *adunea-m*
with beloved NEG was.able-1SG DMS REFL met-1SG
'I was unable to meet with my sweetheart'

and explains it as the process from Albanian ("le procédé de l'albanais") manifesting itself in Aromanian. Still, this tense-matching appears to be far more extensive in Albanian than in Aromanian.

This requirement of tense-matching affects the complement semantics in the sense that the imperfect form, e.g. *qëndronte* in (25), does not have its usual interpretation as a past tense progressive aspect form; rather, it has a special meaning in conjunction with *të* and a past tense matrix verb. It is as if the use of just the subjunctive here in a past tense form is elliptical for '(decided) that he would stay' (*ai vendosi se do të qëndronte*) with an indicative complementizer *se* and the conditional form of the verb, consisting of the future marker *do* plus the DMS *të* with the imperfect tense form *qëndronte*, meaning 'he would stay'. However, the pattern, even if elliptical in some sense, perhaps in origin, has taken on a value as a grammatical requirement for at least some styles and some speakers; note too that the Aromanian example does not lend itself to an elliptical interpretation, suggesting a grammatical value there, with language contact, via calquing, at its core.

5 Beyond complementation

Although the focus here has been on complementation in the strict sense of clauses that fill sentential argument requirements, and on the complementizers and markers, or the absence thereof, that signal them, there are related uses of

these elements that show their versatility and semantic range. Two key aspects in this regard are treated here.

DMS clauses, as argued in Section 2, are mood-marked clauses with zero (null) complementizers. While they are generally restricted to occurring thus in clauses that are subordinated to main clauses, DMS clauses – the same clauses that figure in modal complementation – can also occur as main clauses and impart a modal sense to an utterance. Since this usage involves the extension into a matrix clause of a verb form typically associated with use in a subordinate clause, it can be called, following Evans (2007), "insubordination".

Newmark, Hubbard and Prifti (1982: 80) describe the meanings associated with such main clause DMS clauses in Albanian as "the modality of possibility, obligation, or desirability... [expressing] the speaker's desire for an action"; moreover, there are nuances of uncertainty and deliberation as well, in addition to some jussivity and hope. Examples include the following.

(27) a. *Ku* ***ta*** *vë-më?*
where **DMS;it** put-1PL
'Where might we put it?'

b. *Kush* ***të*** *je-të* *kaq* *vonë?*
who **DMS** be-3SG.SBJV so late
'Who can it be this late?'

c. *Ç'* *fjalë* ***t'*** *i* *shpie* *t-im* *at-i?*
what word **DMS** him.DAT lead.1SG DAT-my father-DAT
'What message am I to take to my father?'

d. ***Të*** *vendos-im!*
DMS decide-1PL
'Let's decide!'

e. ***Të*** *rro-jë* *Parti=a*
DMS live-3SG.SBJV Party=ART
'May the Party live on!'

As the examples in (27) show, the semantic range of such insubordinate uses in Albanian is quite extensive, covering various kinds of modal force. Similar examples can be found in the other languages, such as those in (28) from Greek and in (29) from Bulgarian.

(28) a. *pu* ***na*** *kaθisu-me?*
where **DMS** sit-1PL
'Where shall we sit?'

b. ***na*** *sas* *zis-i* *to* *moro*
DMS you.DAT live-3SG the.NOM baby.NOM
'May (your) baby live (long) for you!'

c. ***na*** *mu* *ðosi-s* *ta* *resta* *amesos*
DMS me.DAT give-2SG the change immediately
'Give me the change at once.'

(29) a. ***Da*** *s-te* *živi* *i* *zdravi!*
DMS be-2PL.PRS alive and healthy
'I wish you many happy years' (literally: "May you be alive and healthy")

b. ***Da*** *ti* *se* *ispăln-jat* *vsički-te* *ti* *želanija!*
DMS you.DAT REFL come.true.PFV-3PL.PRS all-ART you.DAT wishes
'May all your wishes come true for you!'

c. ***Da*** *na-piše-š* *tova* *pismo!*
DMS PFV-write-2SG.PRS this letter
'You should really write this letter!'

d. ***Da*** *za-tvori-š* *vrata-ta!*
DMS PFV-close-2SG.PRS door-ART
'Close the door!' (literally: "You should/might close the door.")

From a structural standpoint, it is noteworthy that there are instances with insubordinate DMS clauses in which a WH-word fills the complementizer position, as in (27a)–(27c) and (28a).

As an aside, it should be noted that insubordination is possible only with the modal complement clauses, not the indicative clauses; that is, there are no standalone clauses with just the indicative complementizer, as (30a) from Albanian and (30b) from Greek (where the glosses are an attempt to give a meaning that such an utterance might have).

(30) a. ****Se*** *u takuam pasdrekë* cf. (20b)[30]
'Had we (only) met in the afternoon!'

b. ****Oti*** *θa nikisi i omaða mas* cf. (1bii)
'(O) that our team will win!'

The facts of insubordination with DMS clauses and those in (30) provide another argument that the DMS, as far as the Balkan languages are concerned, is not a complementizer (see Section 2) but rather a mood marker, since its main-clause

30 There is one interesting example in Newmark, Hubbard & Prifti (1982: 84) that has a clause headed by the invariant complementizer *që* (seen above in [5a] and [14]) occurring independently:

(iii) *Po,* ***që*** ***t'*** *I* *ke-të* *vajtur* *mendja* *asaj* *apo* *mua*
but **COMP** **DMS** him.DAT has-3SG.SBJV gone mind him.DAT or me.DAT
se ... kurrë
that never
'But, that it may have crossed her mind or mine that ... Never!'

This may be elliptical for "the idea that ..." or the like, but it suggests an insubordination for at least some non-modal complementizer clauses. This merits further investigation, to be sure.

behavior is different from that of the canonical complementizers like Albanian *se* and Greek *oti*.[31]

Once one moves away from complementation per se, and looks at subordinate clauses more generally (even those used insubordinately), as the examples in (22) and (23) at the end of Section 3 indicate, one finds subjunctive clauses with the Albanian *të* occurring far more widely in combinations with other subordinating elements, even if they do not function as argument complements. A few examples are given in (31), including some that are like (27) with WH-words but in this case, embedded.

(31) a. *po* ***t'*** *i* *zgurdullo-sh* *sytë*
if **DMS** him.DAT pop.open-2SG.SBJV eyes
'if you make his eyes pop out ...'

b. *edhe* *sikur* ***ta*** *vi-nin* *re*
even as.if **DMS;it.ACC** put-3PL.IPFV novelty
'Even if they noticed it ...'

c. *pasi* ***të*** *na* *ke-në* *lënë*
after **DMS** us.ACC have-3PL.SBJV left
'after they have left us ...'

d. *Këtu* *ka* *qiell* *sa* ***të*** *dua-sh*
here has.3SG sky however.much **DMS** want-2SG
'Here you have however much sky you may want!'

Example (31a) is important, as it brings to light a way in which Albanian differs from Greek. In particular, Albanian allows the DMS to occur with the conditional subordinator *po* 'if'; Greek here has only the pattern in (32a) without the DMS and never allows the pattern in (32b) with the DMS:

(32) a. *an* *kani-s* *fasaria, ...*
if make-2SG fuss
'if you make a fuss, ...'

b. **an* *na* *kani-s* *fasaria*
if DMS make-2SG fuss
'if you make a fuss, ...'

Moreover, the Greek verb with *an* 'if' is demonstrably not subjunctive, because the negator used is the indicative negator *ðen* and not the modal negator *min* (cf. *an ðen kanis* / **an min kanis* 'if you do not do').[32] However, when 'if' is used as a

31 There are languages in which insubordination is found with constructions involving undeniably canonical complementizers (e.g. French and Spanish with their respective realizations of *que*) so that this behavior would provide a Balkan-specific argument only.

32 Evidence for *mi(n)* as a modal negator is the fact that only it can negate a verb marked with the DMS *na* and *na, so* that (iv) is perfectly well-formed.

canonical complementizer, for which the translation 'whether' is appropriate, a different Albanian form is used, *në* or *nëse*,[33] and *të* is excluded:

(33) *S'* *ësht-të* *puna* [***nëse*** *e* *njoh* *apo* *jo*]
NEG be-3SG matter **whether** him know.1SG or not
'It is not a matter of whether I know him or not'

Greek here uses the same form as the conditional subordinator in (32), *an*:[34]

(34) *ðen* *ksero* [***an*** *exi-s* *ðikio* *i* *oxi*]
NEG know-1SG **whether** have-2SG right or not
'I don't know if you are right or not'.

Thus for all the parallelism in the complementation and subordination systems of different Balkan languages, there are differences in detail as well that are equally noteworthy, all the more so in the face of the convergences.

6 Concluding remarks: The role of language contact

This chapter begins in Section 1 with a consideration of how complementation fits into the complex of structural properties that define the Balkan Sprachbund, a construct born of language contact. Thus, in closing this chapter, it is fitting to

(iv) ***na*** ***mi*** *me* *kitaz-is* *etsi*
DMS **NEG** me.ACC look.AT-2SG thus
'You should not be looking at me in that way'

Using the indicative negator *ðe(n)* in that context is impossible: **na ðe me kitazis etsi*. On the modal negator in Greek, see Janda & Joseph (1999), and with a comparative look at the Greek modal negator and its corresponding modal negator in Albanian, see Joseph (2002). Only Romani among the other Balkan languages shows an indicative/modal distinction in negation.

33 *Nëse*, of course, appears to contain the indicative complementizer.

34 This example highlights an interesting property of the Greek verb *ksero* 'know': besides occurring with the conditional complementizer, it can also occur with a modal (DMS) complement with slightly different semantics, meaning 'know how', and with an indicative complement in the meaning 'know that', e.g.:

(v) *kser-o* [***na*** *kolimbis-o*]
know-1SG **DMS** swim-1SG
'I know how to swim'

(vi) *kser-o* [***oti*** *i-se* *kalos* *anθropos*]
know-1SG **COMP** are-2SG good man
'I know that you are a good man'.

turn once again to a further consideration of the role of language contact in complementation in the Balkans, by way of seeing how language contact is responsible, or not, as the case may be, for the facts discussed herein. Besides shedding light on the Balkans per se, this final section offers a view of how complementation fares in situations of language contact.

One key point about language contact and complementation in the Balkans is that Balkan complementizer words themselves show evidence of borrowing. This is an interesting development because complementizers are function words, part of a closed class of grammatical items. As such, they are a kind of word-class that is often held to be among the less-easily borrowed items, like pronouns or adpositions or low numerals.[35] Nonetheless, despite this status, Matras (2009: 196) offers many well-documented instances of the borrowing of complementizers, which he attributes to their use in discourse; see also Matras and Tenser (this volume) for more examples, all involving Romani, and Friedman and Joseph (2017: §4.3.3.4) for other cases from the Balkans with additional discussion. And, indeed, in the Balkans, the borrowing of such elements in a variety of functions is widespread, with examples including temporal (35a), causal (35b-d), conditional (35e), and concessive (35f) subordinators. The Bulgarian examples in (35) reflect now-obsolete usage that was more current during the period of the Ottoman Empire, but the fact of the borrowing remains even if the loans are now obsolete. The concessives in (f) are based on Greek *makari*, from Ancient Greek *makar* 'blessed', used in later Greek to mean 'God willing' and thus to serve complementizer-like functions introducing wishes, augmented with native subordinators.

(35) a. Ayia Varvara Romani *molis* 'as soon as' < Greek *molis* 'as soon as' (Igla 1996: s.v.)
b. Bulgarian *zerem* 'because' < Turkish *zira* 'because' (Grannes, Rå Hauge & Süleymenoğlu 2002: s.v.)
c. Bulgarian *čunki(m)* 'because' < Balkan Turkish *çünki* 'because'[36] (Grannes, Rå Hauge & Süleymenoğlu 2002: s.v).
d. Bulgarian and Macedonian *oti* 'because, for that reason' < Byzantine Greek *óti* '(for) that which'
e. Aromanian *ama că* 'if'' < Greek *áma* 'when, if' (with Aromanian subordinator)
f. Aromanian *macar(im) si* 'even if'
Bulgarian *makar če* 'even though', *makar i da* 'even if; although'
Macedonian *makar što* 'even though', *makar i da* for 'even if; although'
Vlax Romani *màkar kę* for 'although', *màkar te* 'even if' (Hancock 1995: 113)

35 See Thomason & Kaufman (1988) for relevant discussion; the now-famous "Swadesh list" of meanings resistant to borrowing and replacement (Swadesh 1950) also includes such words, and is thus an early statement of this view.

36 The contemporary Standard Turkish form is *çünkü*, but the form as given reflects Balkan Turkish phonology.

Moreover, throughout Balkan Romani, as discussed more fully in Matras and Tenser (this volume), although there is a native form, *kaj* (from earlier 'where?') that serves as an indicative complementizer, it is also the case that forms borrowed from the co-territorial dominant language are often used; thus in the Romani of Greece *oti* 'that' occurs, from Greek *oti* 'that', while in the Romani of Bulgaria *či* 'that' occurs, from Bulgarian *če*, and in the Romani of Romania *ke* occurs, from Daco-Romanian *că*. The occurrence of borrowed indicative complementizers in Romani correlates with a grammatical effect that goes beyond the mere addition of a complementizer lexeme to the language. In particular, the borrowed indicative complementizer comes to occur alongside the (native) DMS *te*, thus giving the language a structural distinction of indicative versus modal complementation just like that found in the other Balkan languages, as discussed earlier. It seems, then, that this Balkan distinction has been carried over into Romani and established through the language contact that gave rise to these indicative complementizer borrowings. It may be, however, that what we see in Romani is the maintaining of an already-existing distinction through the borrowing; that is, this distinction might have predated the entry of Romani into the Balkans.[37] Still, the situation is suggestive of the potential that intense speaker contact with other languages can have to reach even into a native language's semantics of complementation and into the network of interrelationships among subordinating elements.

In addition, there is calquing of complex complementizers in which speakers of one language produce a morpheme-by-morpheme loan translation of a composite form in another language. For example, Aromanian has the composite forms *s-easte că* / *s-fúre că* for 'if', both based on forms of the verb 'be'. Inasmuch as *fúre* is from *fuerit*, the Latin perfect subjunctive of 'be', *s-fúre că* looks rather like Albanian *në qoftë se* 'if; in case that', literally "in may-it-be (optative) that", so that the similarity can be accounted for by positing a loan translation by speakers of the one language, calquing the form of the other language; what is uncertain here, however, is the directionality of the loan translation, specifically who calqued from whom.

In some instances, the calquing leads to parallelism in usage. Again without any clear indication as to the directionality, one finds in Greek, Bulgarian, and Macedonian, in the standard languages, in varieties of Albanian and Romani,

37 Some equivocation is needed here because it is possible that the Balkan Romani distinction continues one made in earlier Indic. Sanskrit, representing older Indic, and various modern Indic languages do not seem to make the indicative/modal distinction via the choice of complementizer, though there are both complementation via infinitives and complementation with a finite verb that offer a somewhat analogous structural distinction to the Balkan one, if not a semantic one.

and in Old Romanian[38] the use of the preposition meaning 'for' (*ja*/*za*/*za*/*për*/*za*/*pentru*, respectively)[39] together with the language-specific DMS and a finite verb in the meaning 'in order that …'; these are illustrated in (36) with the first person singular of the verb 'write', thus 'in order that I write / in order for me to write'.

(36) a. *për të shkruaj* (Albanian)
b. *za da piša* (Bulgarian)
c. *ja na γrapso* (Greek)
d. *za da pišam* (Macedonian)
e. *pentru să scriu`* (Old Romanian)
f. *za te čhinav* (Romani, Goli Cigani dialect, Bulgaria [Matras & Tenser, this volume])

In each case in (36), the formation is synchronically somewhat anomalous in having a DMS-headed verb ostensibly governed by a preposition, a word that in principle looks for a noun phrase, not a verb, as its complement.

As a final point about language contact and complementation, it is important to realize that not all convergence is due to contact. Languages can show convergent features due to common inheritance from a proto-language, if they are members of the same language family. However, besides contact and inheritance, there can be chance convergences between languages, i.e. the result of completely independent developments in each language, as well as similarities due to universalities, traits that recur cross-linguistically that are due to general properties of human language or the ways in which humans interact with one another or with the world at large.

All this being the case, it is worth considering what aspects of Balkan complementation might be due to factors other than language contact. For instance, the very presence of an indicative-versus-modal distinction itself is not restricted to the Balkans: most Indo-European languages have, or have had, the same or a similar distinction in modality – it is widespread across Romance, was a robust part of earlier Germanic languages, is found in Indic languages, and so on. Thus

38 Old Romanian is given here because Modern Daco-Romanian uses *pentru* with *ca* 'as' followed by the subjunctive with the DMS *să*, or *pentru* with the infinitive (thus, *pentru ca să scriu* 'in order that I write' (literally, "for as that I.write") or *pentru a scrie* 'in order to write' (literally, "for INFM write.INF"), thus with a syntax that is a bit different from the older, more Balkan, construction in (36). Aromanian differs here in having simply *ca să* (literally, "as DMS") for 'in order that'.

39 Strictly speaking, the Romani form in (36f) does not really involve a "preposition" *za*, as it is a borrowed item (from Bulgarian) that seems to be used just in this context; this represents a mix of a calque on the structure of Bulgarian *za da* and an outright borrowing of *za*. It is interesting that *za da te* also occurs in this dialect, suggesting that *za da* was taken to be a unit, but one in need of overt DMS modality.

it could simply be present in Balkan languages by virtue of their Indo-European legacy. However, this distinction is also found in Turkish, presumably an inheritance from Proto-Turkic as it is found in other Turkic languages, such as Uzbek, and in such diverse languages as Cree, an Algonquian language of Canada, and Arabic, a Semitic language. The Balkan convergence of having such a distinction could just as well constitute a universal aspect of human language, reflecting a need to comment on realia, for which indicative would be appropriate, and on irrealis conditions, for which modal would be appropriate.[40] However, the particular formal manifestation of the distinction, in the Balkan case via the choice of complementizer and the use of verbal markers, as opposed to marking via suffixes in the Turkic case, and via a combination of prefixes and special verbal endings in the Cree and Arabic cases, would not be universal. As such, it thus could well be affected by language contact, and could show convergence due to contact.

In closing, it can be noted, in light of this last point, that even in the unlikely event that language contact has played no role in Balkan complementation, the similarities seen among the various languages discussed here would remain as interesting from a typological standpoint. Contact, however, is very likely to have been the source of the similarities, so the intrinsic linguistic interest of the region with regard to the semantics of complementation is thereby enhanced. That is, there is not just a typological motivation for examining complementation in the Balkans, but there is the added dimension of language contact to take into consideration in the study of this aspect of meaning cross-linguistically.

References

Altimari, Francesco. 2011. Traces d'infinitifs anciens appartenant à l'albanais dans la langue arbërisht d'Italie. In Hélène Antoniadis-Bibicou, Maurice Aymard & André Guillou (eds.), *L'homme et son environnement dans le Sud-Est européen. Actes de Xe Congrès de l'Association internationale du Sud-Est européen (AIESEE). Paris, 24–26 septembre 2009* (Textes documents études sur le monde byzantin néohellénique et balkanique 11), 445–455. Paris: Éditions de l'Association Pierre Belon.

40 There are some nonconvergent aspects of Balkan complementation that could be due to universality. For instance, Albanian has an infinitive, the type with *për të* as the infinitival marker, that derives from a purpose expression, being originally "for the-VERB-ing". Given that cross-linguistically, purpose expressions are the most frequent source of infinitives (Haspelmath 1989), the occurrence of such an infinitive in Albanian could be an instantiation of this typologically common strategy.

Camaj, Martin. 1984. *Albanian grammar with exercises, chrestomathy and glossaries*. Collaborated on and translated by Leonard Fox. Wiesbaden: Otto Harrassowitz.
Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.
Chomsky, Noam. 1982. *Some concepts and consequences of the Theory of Government and Binding* (Linguistic Inquiry Monograph 6). Cambridge: MIT Press.
Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness*. Oxford: Oxford University Press.
Friedman, Victor A. 2006. Balkans as a linguistic area. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., vol. 1, 657–672. Amsterdam: Elsevier.
Friedman, Victor A. & Brian D. Joseph. 2017. *The Balkan languages*. Cambridge: Cambridge University Press.
Grannes, Alf, Kjetil Rå Hauge & Hayriye Süleymenoğlu. 2002. *A dictionary of Turkisms in Bulgarian*. Oslo: Novus.
Hancock, Ian. 1995. *A Handbook of Vlax Romani*. Columbus: Slavica Publishers.
Haspelmath, Martin. 1989. From purposive to infinitive – a universal path of grammaticalization. *Folia Linguistica Historica* 10. 287–310.
Igla, Birgit. 1996. *Das Romani von Ajia Varvara: Deskriptive und historisch-vergleichende Darstellung eines Zigeunerdialekts* (Osteuropa-Institut der Freien Universität Berlin Balkanologische Veröffentlichungen 29). Wiesbaden: Otto Harrassowitz.
Janda, Richard D. & Brian D. Joseph. 1999. The Modern Greek negator *μη*(*ν*)(-) as a morphological constellation. In Georgios Babiniotis (ed.), *Greek linguistics: Proceedings of the 3rd International Conference on Greek Linguistics*, 341–351. Athens: Elinika Gramata.
Joseph, Brian D. 2003. The Balkan Languages. In William J. Frawley (ed.), *International encyclopedia of linguistics*, vol. 1, 194–196. Oxford: Oxford University Press.
Joseph, Brian D. 2002. Balkan insights into the syntax of **me*: in Indo-European. In Mark Southern (ed.), *Indo-European perspectives* (Journal of Indo-European Studies Monograph Series 43), 103–120. Washington, DC: Institute for the Study of Man.
Joseph, Brian D. 2010. Language contact in the Balkans. In Raymond Hickey (ed.), *The handbook of language contact*, 618–633. Oxford: Blackwell Publishers.
Joseph, Brian D. 2012. Taming – and naming – the Greek "subjunctive". In Zoe Gavriilidou, Angeliki Efthymiou, Evangelia Thomadaki & Penelope Kambakis-Vougiouklis (eds.), *Proceedings of the 10th International Conference of Greek Linguistics*, 303–308. Komotini: Democritus University of Thrace.
Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
Matras, Yaron & Anton Tenser, this volume. Complementizers in Romani.
Newmark, Leonard, Philip Hubbard & Peter Prifti. 1982. *Standard Albanian. A reference grammar for students*. Stanford: Stanford University Press.
Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description,* vol. II: *Complex constructions*, 2nd edn., 52–150. Cambridge: Cambridge University Press.
Pană Dindelegan, Gabriela. 2013. *The grammar of Romanian*. Oxford: Oxford University Press.
Ross, John R. 1969. Auxiliaries as main verbs. *Studies in Philosophical Linguistics* 1. 77–102.
Sampanis, Konstandinos. 2011. *A diachronic and typological approach to the Modern Greek subjunctive complementation*. Salzburg: University of Salzburg Ph.D. dissertation.

Sampanis, Konstandinos. 2013. Summary of "A diachronic and typological approach to the Modern Greek subjunctive complementation" (2011 University of Salzburg dissertation). *Journal of Greek Linguistics* 13. 165–169.

Sandfeld, Kristian. 1930. *Linguistique balkanique. Problèmes et résultats.* Paris: Champion.

Swadesh, Morris. 1950. Salish internal relationships. *International Journal of American Linguistics* 16. 157–167.

Thomason, Sarah G. & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California.

Vrabie, Emil. 2000. *An English-Aromanian (Macedo-Romanian) dictionary: with two introductory sketches on Aromanian*. University, MS: Romance Monographs.

Natalia Serdobolskaya

Semantics of complementation in Ossetic

1 Introduction

This paper presents an analysis of the system of complement markers in Ossetic (Iranian; Indo-European), a language of the Northern Caucasus. The semantics of non-finite forms and subordinating conjunctions in Ossetic have previously been described in reference grammars and works dedicated to subordination (cf. Abaev 1950; Kulaev 1959; Gagkaev 1956; Bagaev 1982). However, most of these subordination constructions have been dealt with insufficiently and in isolation from each other. In particular, the factors that influence the presence or absence of the correlative pronoun/adverb and the competition between various complement subordinators remain underdescribed.

I consider the different types of complementation in Ossetic in the light of certain semantic parameters that have been elaborated in the research on complementation over the last few decades, such as the oppositions 'fact vs. event vs. proposition' and 'presupposition vs. assertion'. On the basis of these parameters, I try to explain the distribution of the complementation strategies found in Ossetic.

The paper is structured as follows. First, I consider the semantic parameters relevant for the choice of complement construction cross-linguistically, and then I provide an analysis of the Ossetic data.[1]

1 The data presented here were collected by elicitation, from the Ossetic National Corpus (http://www.ossetic-studies.org/iron-corpus/search/?interface_language=en), and from texts recorded during fieldwork. The elicited data were collected in the years 2010–2012 from speakers of Iron Ossetic from Vladikavkaz, Alagir, and Ardon, in the course of work on the "Corpus Linguistics" project of the Presidium of the Russian Academy of Sciences and on RFH projects no. 13-04-00342 and 14-04-00580. I would like to thank Madina Darchieva and Fatima Aguzarova, whose help was invaluable for double-checking the collected material. Examples elicited during fieldwork are given without reference, while examples taken from recorded oral texts are labelled TEXT and examples from the Ossetic National Corpus are labelled ONC.

2 Semantics of complement constructions in the languages of the world

2.1 Definition of complement constructions

A complement construction can be defined by either semantic or syntactic criteria. An example of a semantic definition can be found in (Noonan 1985: 52), where complementation is viewed as a syntactic construction in which "a notional sentence or predication is an argument of a predicate". This means that the semantics is used as the main criterion for the identification of complement clauses. On the other hand, Dixon and Aikhenvald (2006: 1) describe complementation in the following way: "certain verbs can take a clause, instead of an NP (noun phrase), as a core argument". This definition is based on the assumption that the notions of "noun phrase" and "clause" are well-defined and distinguishable in the given language. Both definitions agree in the majority of cases, but give different results for disputable constructions. Consider the following examples:

(1) *He was angry about* ***getting old and getting sick****.*

(2) *He's already angry* ***that I refused to move back in****.*

(3) *He was angry about* ***a war that had taken a child's mother****.*

(4) *I was still angry because of* ***Jilly****.* (COCA)

In (1)–(4) the bold font marks different syntactic constructions used to encode the stimulus of "angry". The semantic definition of complementation covers (1) and (2), which is expected, but also (3) and (4), since both of them refer to situations, and not to entities. In (3) a non-derived noun is used to refer to a situation, while in (4) the NP *Jilly* refers to some situation by association: obviously, it is not the existence of the particular person that has caused the speaker's anger, but some action that she has performed. Classifying (3) and (4) as complementation, runs again most current typological and theoretical work on this issue, and is thus unwelcome.

The syntactic definition gives the predicted results for (1)–(4), rejecting the latter two. However, it excludes many constructions that are in complementary distribution with indisputable "complement constructions", e.g.:

(5) *I like* ***it when everyone is smiling****.* (COCA)

(6) *I love* ***the way she dances****.*

The dependent clause in (5) is introduced by the temporal subordinator *when*, whereas the direct object position of the complement-taking verb is filled by the "dummy object" *it*. Clearly the dependent clause is an argument of the complement-taking verb in semantic terms, but it is less clear whether the sentence has a similar structure at the level of syntax, or whether the dummy pronoun should instead be analyzed as the direct object. (Note that in Ossetic the literal translation of (5) is a complement construction, cf. Section 3.3.2.)

From syntactic point of view, the construction in (6) consists of a relative clause adjoined to the NP *the way*. Such contexts are not considered under the heading of "complementation" in most works on clausal complements (cf. Noonan 1985; Dixon & Aikhenvald 2006: 19). However, semantically this is indeed a complement construction, since the situation *she dances* fills the valency slot of stimulus associated with the verb *love*. Thus (6) can be paraphrased by sentences like *I love her dancing* or *I love how she dances*, both unquestionable examples of a complement construction. In a large number of languages (e.g. many Altaic and Uralic languages) the only way to express manner in complement clauses is by means of nominalizations, including specific verbal nouns denoting manner. These are analyzed as complement clauses in grammars and specialized studies on the languages in question, while their English translation (6) is considered to be a relative clause. However, if we are aiming to present a detailed account of the semantic distribution of clausal constructions expressing stimulus with the verbs *like* and *love* in English, cases like (5) and (6) should not be excluded from our analysis.

According to the syntactic definition, complement constructions headed by nouns should not be described as complement clauses. The problem is that this excludes not only English constructions of the types (5) and (6), but also the Japanese construction with *koto* "thing, fact" together with similar constructions elsewhere. This is hardly satisfactory, as the *koto*-clause is the main complementation strategy used with some complement-taking predicates, and it is usually included in works on complementation in Japanese (cf. Josephs 1976; Suzuki 2000). Hence, the purely syntactic definition of sentential complementation seems to be too exclusive.

For a number of languages we possess syntactic tests to distinguish between complement and adverbial or relative clauses: e.g. Ross's (1967) island constraints filter out (6) (but not [5]). However, these constraints do not work in the same way in all languages. For example, in Ossetic they distinguish between factive and non-factive complement clauses. However, Ossetic has another, highly specific diagnostic: the type and the possibility of omission of the correlative pronoun in the subordinate structure indicate whether it is a complement or an adverbial clause (cf. Section 3.3.2 and Belyaev & Serdobolskaya, in project, for details).

My approach is to combine the two definitions, while making use of language-specific tests. Therefore, I treat as complementation those constructions where one of the semantic arguments of the verb denotes a situation *and* has clausal structure; if language-specific tests exist, they are used as a filter. I do not consider constructions with phasal, aspectual and modal predicates in detail here, since they show monoclausal properties in Ossetic, as is common typologically (cf. Aissen 1974; Noonan 1985). I use the term "complementizer" for lexical/morphosyntactic devices whose main function is that of complementation, and "complementation strategy" for the whole construction of the complement clause (e.g. parataxis, combinations of pronouns, complementizers and special mood on the subordinate verb).

In what follows I will use the following terms: "complement-taking predicate", or CTP – the predicate that can take clausal (and potentially also nominal) arguments; "matrix clause" – the clause with the complement-taking predicate (CTP); "complement clause" – the clausal argument (marked with square brackets); "dependent/embedded clause" – any type of subordinate clause.

The list of matrix verbs analyzed here for Ossetic is given in the Appendix.

2.2 Classes of CTPs and semantics of clausal complements

Many general works on complementation or treatments of this issue in reference grammars present a variety of complementation systems distributed over the CTPs found. The predicates that can take sentential complements are enumerated, and examples illustrating the complementation devices that can be used with each predicate are provided. The description is thus structured as follows: verbs with the meaning "begin", "end", "be able to" take infinitival complements, "see" and "hear" take one type of complementizer and the verbs "think" and "say" take another type of complementizer, and so on. Groups of predicates (e.g. mental, perception verbs, etc.) are proposed on the basis of the devices used for marking their complements. For example, Givón (1980) proposes a hierarchy of CTPs, arguing that the type of CTP determines the verb's choice of complementation device (cf. also the classification of CTPs in the typological works Nedjalkov 1979 and Xrakovskij 1985).

Such classifications, however, find it difficult to account for cases where one and the same CTP can take more than one complementation strategy with different semantics, e.g.:

(7) a. *I like* [*her singing*].
b. *I like* [*how she sings*].

c. *I like* [*that she sings well*].
d. *I like* [*to wake up early in the morning*].

As can be seen from (7), the verb 'like' can take at least four different complementation strategies in English. The complementation strategy chosen depends on the semantics of the complement clause. Determining this verb's position in the hierarchy is therefore problematic. It can be assumed that with each complementizer in (7) the verb 'like' is being used with a different meaning (or nuance of meaning). In this case, our classification would be forced to posit four different verbs: "like 1", "like 2", "like 3", "like 4". Such a treatment may be more or less suitable for distinguishing (7a)–(7c) from (7d), since in (7d) *like* shows a semantic shift away from pure emotion towards decision: while (7d) can be paraphrased as 'I choose to wake up early, but I don't enjoy it', *like* in (7a)–(7c) is semantically close to *enjoy* (cf. the opposition of "direct interaction" vs. "primary consciousness" or "indirect interaction" in Verspoor 2000). However, the difference between (7a) and (7b) and even between (7a)–(7b) and (7c) seems too subtle to be easily explained.

Predicates with similar semantics can display different polysemy patterns cross-linguistically, with the result that they also take different complementation strategies. For example, in many Nakh-Daghestanian languages the verb "want" also means "love"; in many Finno-Ugric languages and in Ossetic the verb "know" also means "be able to"; in Ossetic there is a verb meaning both "think" and "remember". Such polysemy makes it possible to use these verbs with a large number of complementation strategies. On the other hand, it has been shown that even with one and the same meaning of the CTP the complement clause can be introduced by many complementation strategies.

The accumulation of data from various language families led to the development of an alternative approach. This approach is based on the assumption that the dependent clause can be described as having its own semantics, separately from the CTP. The choice of complementation strategy encodes the semantics of the complement clause (Ransom 1986; Podlesskaya 1990; Dixon & Aikhenvald 2006; Serdobolskaya 2009; Serdobolskaya, Motlokhov 2009). The semantics of the dependent clause can be encoded by the complementizer, the mood of the dependent verb, particles, special correlative pronouns etc. The terms "fact", "event", "proposition", "presupposition", "assertion", "truth value", "modality" and "epistemic value" have been drawn on to explain the distribution of the alternative complementizers used with one and the same CTP. For example, (7a) is described as an "event", "occurrence", or "instantiation of the situation", (7b) as a manner complement clause, (7c) as a factive complement clause, (7d) as "action" or "potential action", etc.

In the next section, I will briefly discuss the notions that will be used in this paper, their definitions, and the diagnostic tests used for distinguishing between the phenomena they designate.

2.3 Definitions of semantic types of complement clauses

Beginning in the 1960s, a number of notions have been elaborated to describe CTPs and the semantics of complement clauses. One is the notion of "implicative verbs" (Karttunen 1971: 349): these are verbs that require that "the illocutionary force of S_1 (i.e. assertion, command, question etc.) is shared by S_2", e.g. *John managed to open the box* implies *John opened the box*, but *John hoped to open the box* does not.

Another useful notion is that of "factivity". Kiparsky and Kiparsky (1971: 348) define the "fact" as the "proposition the speaker presupposes to be true". A classic test for the presupposed status of the complement clause is the scope of negation:

(8) *It is odd* [*that the door is closed*]. *It is not odd* [*that the door is closed*].
→ The proposition "The door is closed" is true (Kiparsky and Kiparsky 1971: 349–351)

In both cases, no matter whether the CTP is affirmative or negative, the truth value of the dependent clause is T (true), since it is presupposed to be true. With non-factive complements there is no such presupposition. For example, in the three sentences in (9) the hearer is not supposed to assume that the complement clause is true, even if s/he acquires some information about Joan's and the speaker's opinions on the situation.

(9) *Joan said* [*that the door was closed*]. *Joan did not say* [*that the door was closed*]. *I do not believe* [*that the door was closed*].
→ the truth of the proposition "The door is closed" is not asserted

Kiparsky and Kiparsky (1971) show that the distinction between facts and non-facts is relevant for the encoding of English complement clauses. Facts can be introduced by the gerund with the genitive, while non-factive complements cannot, cf.:

(10) a. *I don't mind* [*your saying so*].
b. * *I maintain* [*your saying so*].
(Kiparsky and Kiparsky 1971: 347)

The expression *do not mind* introduces facts, and the verb *maintain* takes non-factive complements; hence, the unacceptability of (10b). This verb can, however, take *that*-clauses introducing either facts or non-factive complements; see (11).

(11) *I maintain* [*that he did this for sheer vanity*].
(COCA)

It is noteworthy that facts can appear with both factive (like *know*) and non-factive predicates (like *say*), which means that the semantics of the complement clause can be defined as factive independently of the CTP.

Many works use the notion of "proposition" as opposed to "fact" in complement clauses. A proposition is defined as a mentally processed situation that has truth value and is not presupposed to be true (cf. Peterson 1997; Dik 1997; cf. propositional contexts with positive factive and factive epistemic verbs in Asher 1993; cf. the notions of 'predetermined' and 'non-predetermined' truth value in Ransom 1986). This means that the proposition belongs to the assertion made in the complement clause, as in (11). The main difference between fact and proposition thus lies in the 'presupposed' vs. 'asserted' status of the complement clause. This difference between fact and proposition is often demonstrated by contrasting complement clauses of verbs of knowing vs. verbs of thinking (cf. *He knows that the Earth is round* and ?*He thinks that the Earth is round* – the strangeness of the second sentence is due to the presentation of a well-known truth as part of the assertion; putting this under the scope of a verb of opinion suggests that it could be a subject to doubt on the part of the interlocutors).

This definition entails that presupposed information cannot be negated by the same speaker in the subsequent context, e.g. *He thinks/*knows that Joan has left, but that is not true*. This is, however, possible with non-factive complements, introduced (for example) by the verb *think*.

Benveniste claims that, by definition, sentential complements of performative verbs cannot be factive (Benveniste 1966: 272). This applies to certain semantic types of CTPs, such as commissives (*promise* etc.), exercitives (*appoint*, *dismiss* etc.) and some others: in their performative use, these verbs require that the situation in the sentential complement is not true until the situation in the matrix clause is realized (e.g. *I declare you husband and wife* in its performative use makes the complement true upon pronunciation of this sentence; until this moment the complement must be false). The standard performative context is first person singular in the present tense (Benveniste 1966). I suggest to distinguish facts from propositions even in less canonical contexts, including the past tense (*He declared them husband and wife*), on the condition that the complement is made true by the fact that the matrix clause is true (e.g. if the context suggests

that the subject of the matrix clause has the authority to make the complement true etc.). I will refer to such contexts as "quasi-performative contexts".

It has been claimed that certain contexts can be "presupposition-opaque" (Krejdlin 1983; Apresyan 1995). See the following examples from English:

(12) a. *If I **knew** [that by cutting off an arm or cutting out my liver I could be rid of you forever], I would seize the knife and relish the pain and loss, all for the sake of freedom* (COCA)
b. *But if I **knew** [that it was going to take off as fast as it did], I – I certainly would have been in some kind of aerobics program six months prior to releasing the thing.* (COCA)

In (12a), the truth of the complement clause is not presupposed: it is merely hypothesized by the speaker, and the sentence could continue "but I knew that it would not help, so I didn't cut off my arm etc.". There is no presupposition failure, since the presupposition is only present in the imaginative world created by the protasis of the conditional. However, this context is not always presupposition-opaque, since with a different intonation the same complement can be presented as true, cf. (12b), where part of the complement (namely, *it did*) explicitly indicates its truth.

Therefore, the following diagnostics can be used to distinguish between facts and propositions:

– facts cannot be negated by the same speaker in the following context;
– facts cannot occur in the complements of performative CTPs with commissive or exercitive meaning.

I also consider that complements introduced in the protasis of conditional sentences but negated by the subsequent context contrast with genuinely factive complements.

Another important distinction considered in works on the semantics of abstract nouns is the opposition between "propositions" and "events" (cf. Asher 1993; Peterson 1997; Arutjunova 1988; Zaliznjak 1990), "truth" and "occurrence" (Ransom 1986), or "facts/possible facts" and "states-of-affairs" (Dik 1997). Facts/propositions have truth value and thus denote situations that have been "mentally processed" by the speaker, while events are situations that have not been mentally processed; for example, the complements of the predicates 'take place', 'happen' and of immediate perception verbs denote events; cf.:

(13) a. [*Fighting*] ***took place** in the neighbourhood.*
b. *I **watched** [her sing/singing].*

A large number of tests for differentiating between facts (or propositions) and events have been suggested in the literature (cf. Asher 1993, Peterson 1997, Arutjunova 1988, Zaliznjak 1990):

- facts/propositions can contain negation, while events cannot[2];
- facts/propositions and events have different identity conditions (if an NP is substituted by another coreferring NP the identity of events is preserved, while the identity of facts is not);
- facts/propositions are not located in space and time (??? *The fact that ... happened yesterday*);
- facts/propositions cannot be perceived directly by the senses;
- facts/propositions do not have duration (??? *The fact that... lasted two weeks*).

Another diagnostic is proposed by Boye (2012): propositions can host epistemic expressions, while events cannot, cf. the unacceptability of ??? *I saw him maybe run(ning)*, ??? *I'm afraid of maybe going to the forest*.

From here on, when characterizing the semantics of the complement clauses, I use the terms 'fact' vs. 'event' vs. 'proposition' as defined above. I also use the notions 'irrealis' and 'generic event'. I define 'generic event' as an event with generic reference (Serdobolskaya 2011); cf. (14b).

(14) a. *I liked your singing (today).*
b. *I like your singing (at any point in time).*

I also use the term 'irrealis' for propositions with irrealis modality (as in *I don't know if John is here*) or for those that bear the truth value 'false' (*I don't believe that John is here*). This is in accordance with the definition of irrealis adopted by Palmer (2001:1) from (Mithun 1999:173): "[t]he realis portrays situations as actualized, as having occurred or actually occurring, knowable through direct perception. The irrealis potrays situations as purely within the realm of thought, knowable only through imagination". The type of irrealis complements embraces the notions of "indeterminate truth" (also "undetermined truth") as defined in (Ransom 1986); cf. also the notion of "complete uncertainty" (Boye 2010a).

The relevance of these notions for complementation has been demonstrated for a number of unrelated languges (cf. Noonan 1985; Peterson 1997; Podlesskaya 1990; Dixon & Aikhenvald 2006; Serdobolskaya 2009; Serdobolskaya & Motlokhov 2009; Serdobolskaya et al. 2012). These notions will be used in order to

2 This test can be used with the following admonition. Complements including negation can in fact be found in the context of immediate perception verbs; however, such examples are very special cases of violation of expectations and require a special pragmatic context. See (Miller 2003) for discussion and semantic analysis.

explain the distribution of complementation strategies. Such an approach can account for the fact that one and the same CTP can take various strategies (as in [7a]–[7d] above): a CTP can be used in various meanings, or in a single meaning which is compatible with dependent clauses of different semantic types. Compare the English verb *see* in (15) and the Ossetic verb *žonən* in (16):

(15) a. *I saw him enter.*
b. *You'll soon see that I'm not mistaken.*
(COCA)

(16) a. *čəžg* *žon-ə* [*wɜlibɜχ-t-ɜ* *kɜn-**ən***]
girl know-PRS.3SG walibax-PL-NOM.PL do-**INF**
'The girl can cook walibaxs'
(Ossetic national flatcakes with cheese).
b. *mɜ žɜrdɜ* *=jɜ* *žətː-a,* [*ɜvdiw* ***kɜj*** *u*]
my heart it.GEN know-PST.3SG evil.spirit **COMP** be.PRS.3SG
'My heart knew that it was an evil spirit.'
(ONC)

The structural difference between (15a) and (15b) is due to the difference in the semantics of the complement clause: in (15a) the complement clause describes an event, while in (15b) it describes a proposition that (unlike an event) can take negation; recall the tests above. The fact that the speaker is not mistaken cannot be seen directly by the hearer, since there is no such negative event in the real world. Hence, these constructions are used with different meanings of the CTP *see*, namely immediate perception in (15a) and "cognitive" perception, i.e. inference on the basis of perception, in (15b). For this peculiarity of verbs of perception see (Noonan 1985: 129 and Boye 2010b).

The examples in (16) illustrate two different meanings of the Ossetic verb *žonən* 'know, be able to': in the meaning 'be able to, know how to' this verb takes the infinitive, whereas in the meaning 'know that' it takes a finite complement with the subordinator *kɜj*. In such cases, the semantics of the dependent clauses is directly related to the meaning of the CTP.

3 Complementation in Ossetic

3.1 Ossetic language: typological features

The Ossetic language (with 493,610 speakers according to the Russian census of 2002) belongs to the Indo-Iranian subgroup of the Indo-European language

family. The principal dialects are Iron (which forms the basis of Standard Ossetic) and Digor. This study is based on Iron data.

Ossetic shows agglutination in its nominal and fusion in its verbal morphology. It exhibits case alignment of the accusative type, is predominantly dependent-marking in NPs, and has SOV as its basic word order (cf. Abaev 1950; Bagaev 1965; Axvlediani 1963).

There are nine nominal cases in Iron Ossetic: nominative, genitive, dative, allative, ablative, inessive-illative, superessive-superlative, equative, and comitative. It is important to specify that the genitive case has two main functions, marking not only the possessor in an NP but also the direct object of a verb. In this latter function the genitive marker can be dropped (exemplifying the typological phenomenon of differential object marking). The distribution of the genitive marker is mostly based on animacy (although information structure and referential properties are also relevant): animate direct objects appear with the genitive marker, while non-animate DOs remain unmarked.

The morphology of the Ossetic verb includes three tenses of the indicative (present, past, and future) and four oblique moods (imperative, subjunctive, optative, and counterfactive) (Vydrin 2011). The main opposition in the aspect system is between imperfective and perfective; the perfective is encoded by prefixes on the verb.

In the domain of subordination, finite clauses with overt subordinators are most often used. One of the important characteristics of the subordination in Ossetic is the predominant use of correlates with all subordination types. All three types of subordinate clause – relative, adverbial, and complement clauses – can be formed in the same way, with a subordinator in the dependent clause and a corresponding distal demonstrative ("correlate") in the matrix clause; cf. (17) with a relative, (18) with an adverbial, and (19) with a complement clause.

(17) [*mɜj* *ražmɜ* ***sə*** ***qug*** *ba-lχɜtː-a*], ***wəj*** *wɜj* *kɜn-ə*
month before **what** **cow** PRVB-buy-PST.3SG **that.DEM** sell do-PRS.3SG
'He is selling the cow that he bought a month ago.'

(18) [***salənmɜ*** *=dən* *mɜ=* *sɜšt* *ɜrtːiv-a*], ***walənmɜ***
as.long.as thee.DAT my eye shine-SBJV.3SG **to.that.time**
dɜ= *χis-ɜn* *binont-ə* *koj* *ba-kɜn*
thy self-DAT family-GEN care PRVB-do[IMP.2SG]
'As long as I'm alive, take care of your family.'
(Gagkaev 1956: 227)

(19) *ɜž* *žon-ən,* [*zawər* *uš* ***kɜj*** *ɜr-χašt-a*], ***wəj***
I know-PRS.1SG Zaur wife **COMP** PRVB-bring-PST.3SG **that.DEM**
'I know that Zaur has married.'

In relative clauses and in the most types of adverbial clauses (apart from purposive and substitutive clauses), the correlate is obligatory. In complement clauses, as well as in purposive and substitutive clauses, the construction without the correlate can be chosen instead (cf. Belyaev & Serdobolskaya, in project, for details).

3.2 Overview of complementation strategies in Ossetic

Ossetic has a large number of devices used in complementation:

– the infinitive in *-ən*, which is used with phasal, modal, emotive, mental, causation, speech and evaluative predicates:

(20) [*gɜdə* *nəχaš* *=dɜr* *=zə* *ba-ftaw-**ən***] *qɜw-ə*
cat word ADD it.IN PRVB-add-**INF** must-PRS.3SG
'Well, lies also must be added' (when telling a story).
(TEXT)

– nominalizations headed by the participle in *-t/-d* (homonymous with the preterite stem of the verb; cf. Abaev 1950):

(21) *lɜp:u* [*kaš* *χɜr**d***] *fɜs-i*
boy porridge eat:**PTCP.PST** PRVB-EXST[3]
'The boy ate up the porridge'
(lit. finished eating).

Nominalizations only rarely occur in Ossetic complement clauses; however, they are acceptable with nearly all CTPs.

– participles in *-gɜ* and *-gɜjɜ* (the latter is the ablative form of the participle in *-gɜ*; cf. Belyaev & Vydrin 2011):

(22) *ɜž* *fet:-on* *de=* *fšəmɜr-ə* *bɜχ-əl* *sɜw-**gɜ-jɜ***
I see-PST.1SG thy brother-GEN horse-SPESPL go-**CVBP-ABL**
'I saw your brother riding a horse.'

The participles are only used with verbs of immediate perception (in their direct sense only; the "cognitive perception" reading of [22] is not possible).

– the subordinators *kɜj*, *kʷəd*, *kʷə*, *sɜmɜj*, *ɜmɜ*, *kɜd*, *səma*, *salənmɜ*:

3 In Ossetic, the preverbal existential verb has the meaning 'end, finish'.

(23) *ɜmɜ* *=dam* *=dɜ* *fɜnd-ə,* [***sɜmɜj*** *=dɜ* *fe-rvɜž-ən*
and CIT thee.GEN[4] want-PRS.3SG **PURP** thee.GEN PRVB-be.liberated-INF
kɜn-on *de=* *ldar-ɜj*]*?*
do-SBJV.1SG thy lord-ABL
'And, you say, you want me to free you from your lord?'
(TEXT)

The subordinators and the infinitive are the devices used most often to introduce complement clauses in Ossetic, and the next section focuses mostly on their distribution.

Complement constructions with subordinators can have correlative pronouns in the matrix clause, as in (19) above which contains the 3rd person pronoun *wəj*.

The subordinators used in complementation can be divided into two groups with regard to their syntactic properties. Subordinators belonging to the first group (*kɜj, kʷə, kʷəd*) are only found in preverbal position (before the complement verb). The subordinators of the second group (*kɜd, sɜmɜj, səma, salənmɜ*) can "float" inside the dependent clause, but most often occur clause-initially. The conjunction *ɜmɜ* does not adhere to any of these groups: it can only appear clause-initially.

– the citation particles *dam* and *žɜʁgɜ* (participle of the verb of speech 'say'):

(24) *adɜm-ə* *tɜrš-ən* *kot:-oj,* [*arʁʷəd* *či* *nɜ* *w-a,*
people-GEN fear-INF do-PST.3PL baptism who NEG be-SBJV.3SG
wəj ***=dam*** *žəndon-ə* *qižɜmar* *kɜn-zɜn,* ***žɜʁ-gɜ***]
that.DEM **CIT** hell-IN torture do-FUT[3SG] **say-CVBP**
'They threaten people: 'Those who are not baptized will suffer tortures in hell'".
(ONC)

Citation particles are most often used with speech verbs, but they may also occur with non-speech CTPs (cf. Vydrin, forthcoming, for details on *žɜʁgɜ*).

Citation particles do not take correlative pronouns. Both citation particles can be combined and repeated several times in one and the same clause; they do not necessarily require the presence of a matrix clause. Therefore, they cannot be analyzed as "pure" complementizers.

– parataxis:

(25) *qʷədə* *kɜn-ən,* [*je=* *mbal-mɜ* *a-səd-i*]
thought do-PRS.1SG his friend-ALL PRVB-go-PST.3SG
(Where is your father?) '– I think he went to his friend.'

4 In Ossetic, the experiencer of the CTPs 'want', *wərnən* 'believe' and *qɜwən* 'must' is marked with the genitive.

The paratactic construction consists of two finite clauses without any morphological or lexical complementizer.

– indirect question strategies:

Constituent questions are formed in Ossetic with wh-words, which occur in preverbal position[5] (their linear position is strictly fixed), cf.:

(26) *a-sə* *lɜp:u* ***či*** *u?*
this-ATTR boy **who** be.PRS.3SG
'Who is this lad?'

The same pattern is preserved in indirect questions, cf.:

(27) *nɜ=* *fəd* *ba-faršt-a,* [*a-sə* *lɜp:u* ***či*** *u*]
our father PRVB-ask-PST.3SG this-ATTR boy **who** be.PRS.3SG
'Father asked who that lad was.'

General questions are marked by means of word order and prosody; for the most part, no special question particles are used. For example, with a different prosody the interrogative sentence (28) could be interpreted as affirmative. The same pattern is used in indirect polar questions (29).

(28) *šiχor* *sɜt:ɜ* *u?*
lunch ready be.PRS.3SG
'Is lunch ready?'

(29) *nɜ=* *fəd* *ba-faršt-a,* [*šiχor* *sɜt:ɜ* *u*]
our father PRVB-ask-PST.3SG lunch ready be.PRS.3SG
'Father asked if the lunch was ready.'

In the next section I analyze the semantics of each complementation strategy in Ossetic. First, I consider finite sentential complements, i.e. the paratactic construction and clauses introduced by subordinators or citation particles. Then I describe the morphosyntactic and semantic properties of non-finite strategies, such as the infinitive, the nominalization, and the participle.

5 By "preverbal position" I mean a strictly fixed position in the preverbal domain, since there are lexical units that can occur between the subordinator and the verb (e.g. the negation particle and some adverbs).

3.3 Finite complementation strategies: subordinators and the paratactic construction

The subordinators able to introduce complement clauses are *kɜj*, *kʷəd*, *kʷə*, *sɜmɜj*, *ɜmɜ*, *kɜd*, *səma*, and *salənmɜ*. The distribution of these subordinators is fairly transparent. The subordinators *kɜj*, *kʷəd*, *kʷə* denote fact (or proposition), event, and generic event, respectively. The subordinator *sɜmɜj* is used to encode propositions with future temporal reference (with respect to the situation in the matrix clause). The subordinator *kɜd* and *salənmɜ* are used only with the CTP 'wait'. The paratactic construction, the conjunction *ɜmɜ* and the subordinator *səma* denote propositions. Examples of all these strategies follow.

3.3.1 The subordinators *kɜj* 'that' and *kʷəd* 'how'

The subordinator *kɜj* can denote facts or propositions with mental, emotive, perception, speech, and evaluation predicates:

(30) *qʷədə* *=jɜ* *kɜn-ən,* [*šara* ***kɜj*** *ba-χgɜt:-on*]. *gom* *-sɜmɜn* *u?*
thought it.GEN do-PRS.1SG shed **COMP** PRVB-close-PST.1SG open why be.PRS.3SG
'I remember that I closed the shed. Why is it open?'

With most of these CTPs it contrasts with the subordinator *kʷəd* used to denote events:

(31) [*ɜrəgon-ɜj* *=iw* ***kʷəd*** *kafəd-əštɜm*], *wəj* *=ma* *qʷədə* *kɜn-ən*
young-ABL ITER **how** dance-PST.1PL that.DEM PRT remembrance do-PRS.1SG
'I remember how we danced when we were young.'

In (31) the verb *qʷədə kɜnən* 'remember' takes an eventive complement: the speaker recollects in detail the feelings and emotions felt in the situation, while in (30) what is recollected is merely the fact of an event having occurred. The same distinction is observed in the next pair of sentences: in (32a) the complement clause is mentally processed (cf. the notion of "consciousness" in Verspoor 2000) and evaluated as having positive effect. In (32b) the positive emotion arises as a result of the situation described by the dependent clause without mental processing of that situation.

(32) a. *mɜ=* *žɜrdɜ-mɜ* *sɜw-ə,* [*də* *χorž* ***kɜj*** *kuš-əš*], *fɜlɜ*
my heart-ALL go-PRS.3SG thou good **COMP** work-PRS.2SG but
[*alə* *bon* *ɜrɜžˇ-ə* ***kɜj*** *kɜn-əš*], *wəj* *mɜ=* *žɜrdɜ-mɜ* *nɜ*
every day late-IN **COMP** do-PRS.2SG that.DEM my heart-ALL NEG
sɜw-ə
go-PRS.3SG
(A boss to his subordinate.) 'I like it that you do good work, but I don't like it that you are often late.'

b. *mɜ=* *žɜrdɜ-mɜ* *sɜw-ə,* [*dɜ* ***kʷəd*** *žar-əš*]
my heart-ALL go-PRS.3SG thou **how** sing-PRS.2SG
'I like your singing.'

As in many other languages, different complementation strategies are used to differentiate between immediate and indirect (cognitive) perception; cf.:

(33) a. [***kʷəd*** *zərtː-at*], *wəj* *fe-qʷəšt-on*
how speak-PST.2PL that.DEM PRVB-hear-PST.1SG
'I heard you talking.'

b. *ɜžˇ* *šəχɜg-t-ɜj* *fe-qʷəšt-on,* [*ɜrɜžˇə* *mašinɜ* ***kɜj***
I neighbour-PL-ABL PRVB-hear-PST.1SG recently car **COMP**
ba-lχɜtː-aj]
PRVB-buy-PST.2SG
'I heard from the neighbours that you recently have bought a car.'

In the immediate perception sense, the CTP takes the eventive type of complement with the subordinator *kʷəd*, see (33a), while in the sense of indirect perception (where the meaning of the verb 'hear' shifts towards 'learn, find out') it takes *kɜj*, see (33b).

Verbs of speech can introduce events, facts, or propositions. Eventive complements take the subordinator *kʷəd* (34), while *kɜj* introduces facts or propositions, as in (35).

(34) *zalinɜ* *zur-ə,* [*je=* *rəgon* *bon-t-ə* ***kʷəd*** *kafəd-i*]
Zalina speak-PRS.3SG her young day-PL-IN **how** dance-PST.3SG
'Zalina says how she danced when she was young (*says that she danced).'

(35) *mɜ=* *fəd* *ra-zərtː-a,* [*ɜrɜžˇə* *nog* *χid* ***kɜj*** *š-arɜšt-oj*]
my father PRVB-speak-PST.3SG recently new bridge **COMP** PRVB-build-PST.3PL
(Father came back from the village, and we asked him about the news. What's new, what did father tell you?) 'Father told us that a new bridge has been built recently.'

The subordinator *kʷəd* can denote manner in complementation, as well as in adverbial clauses (36). For example, (32b) can be interpreted as 'I like how you sing'.

(36) *wəj* *kaf-gɜ* *=dɜr* *aftɜ* *χorž* *kɜn-ə,* [*žar-gɜ* ***kʷəd***
that.DEM dance-CVBP PRT so good do-PRS.3SG sing-CVBP **how**
kɜn-ə], *aftɜ*
do-PRS.3SG so
'He dances as good as he sings.'

Another meaning of the subordinator *kʷəd* is that of strict causation, see 3.3.3.

3.3.2 The subordinator *kʷə* 'if, when'

The subordinator *kʷə* encodes generic events with emotive and evaluation predicates, cf.:

(37) a. *ɜž* *warž-ən,* [*χud-gɜ* ***kʷə*** *fɜ-kɜn-əš*], ***wəj***
I love-PRS.1SG laugh-CVBP **when** PRVB-do-PRS.2SG **that.DEM**
'I love it when you laugh.'

The sentence in (37a) presents an event, not a proposition, since it declines the diagnostics of epistemic expressions: no epistemic expressions are acceptable in the complement clause:

(37) b. **ɜž* *warž-ən,* [*χud-gɜ* *ɜnɜmɜng /* *ɜvɜcːɜgɜn* ***kʷə***
I love-PRS.1SG laugh-CVBP undoubtedly possibly **when**
fɜ-kɜn-əš], *wəj*
PRVB-do-PRS.2SG that.DEM
Intended: 'I love it when you undoubtedly/possibly laugh.'

The subordinator *kʷə* is mostly used in conditional and temporal adverbial clauses; it could thus be suggested that (37a) is an example of an adverbial clause. However, in Ossetic there is a clear syntactic evidence for the interpretation of such constructions as sentential complements, and not as adverbial clauses. This evidence is provided by the type of correlative pronoun found in the matrix clause. Complement clauses take the correlative pronoun *wəj* 'that/he/she/it', while adverbial clauses take demonstrative adverbs:

(37) c. ?[*χud-gɜ* ***kʷə*** *fɜ-kɜn-əš*], *ɜž* *=dɜ* ***wɜd*** *fɜ-warž-ən*
laugh-CVBP **when** PRVB-do-PRS.2SG I thee.GEN **then** PRVB-love-PRS.1SG
'?I love you when you laugh.'

With the adverb *wɜd*, the dependent clause is taken to be adverbial, with the result that it becomes difficult for native speakers to interpret (37c) (cf. the translation).

There are emotive verbs that can take both correlative pronouns and adverbs with no apparent difference in meaning:

(38) [*birɜʁ* ***kʷə*** *fɜ-waš-ə*], ***wɜd*** / ***wəm-ɜj*** *tɜrš-gɜ*
wolf **when** PRVB-howl-PRS.3PL **then** **that.DEM-ABL** fear-CVBP
fɜ-kɜn-ən
PRVB-do-PRS.1SG
'When the wolves howl, I get scared / I'm scared of the howling of wolves.'

The two constructions in (38) are differentiated syntactically: the correlate *wɜd* introduces adverbial dependent clauses, while *wəmɜj* introduces complement clauses.

Another function of the subordinator *kʷə*, observed with a small number of CTPs (e.g. 'want', 'wait', marginally with *χorž* 'good'), is the encoding of propositions that fill the valency slot assigned to the stimulus:

(39) *me=* *mžɜχχon-t-ə* *fɜnd-ə*, [***kʷə*** *=šən* *ba-χχʷəš*
my compatriot-PL-GEN want-PRS.3SG **if** they.DAT PRVB-help
kɜn-iš], *wəj*
do-OPT.2SG that.DEM
'My countrymen want you to help them.' (ONC)

With these verbs this subordinator can only be used to encode a (non-)desired situation. With *χorž* 'good' it occurs only marginally, and encodes desire ('It would be good if...') rather than pure evaluation.

3.3.3 The *subordinators sɜmɜj* and *kʷəd* 'in order that'

The subordinator *sɜmɜj* 'in order that' is used to mark complements with future reference (with respect to the temporal reference of the matrix clause), (40), or with gnomic meaning, (41b).

(40) *aχʷərgɜnɜg* *žaχt-a*, [***sɜmɜj*** *šk'ola-mɜ* *ɜrba-sɜw-at*]
teacher say-PST.3SG **PURP** school-ALL PRVB-go-SBJV.2SG
(A boy says to his parents:) 'The teacher said that you should come to school.'

Most CTPs that take this subordinator require future reference in their complement; however, not all of them share this requirement, see, for example, 'love' in (41a). With many CTPs the subordinator *sɜmɜj* competes with the infinitive. Roughly speaking, the infinitive is only possible in control contexts, where the semantic subject of the infinitive is coreferential with the subject/object of the CTP, while the subordinator can be used irrespective of the coreferentiality pattern. However, the subordinator is not acceptable with many verbs that take the infinitive, such as 'try', 'get used to', 'teach', 'promise' (see the Appendix). The infinitive can encode both events and propositions (with future reference or

gnomic meaning), while the subordinator has narrow semantics and can only encode propositions with future reference or gnomic meaning:

(41) a. *ɜž warž-ən* [*kaf-**ən***]
I love-PRS.1SG dance-**INF**
'I love dancing.'
b. *ɜž warž-ən,* [***sɜmɜj*** *rɜšuʁd w-a*]
I love-PRS.1SG **PURP** beautiful be-SBJV.3SG
'I love it to be beautiful.'

For example, the verb 'love' takes the infinitive in (41a), where the speaker describes his/her emotions felt in the situation of dancing, while in (41b) the situation in the complement is being evaluated as positive. However, with the verb 'love' examples making use of *sɜmɜj* are attested rarely (the subordinators *k^{w}ə*, *k^{w}əd*, and *ɜmɜ* are much more frequent in the corpus).

The verb 'say' takes the subordinator *sɜmɜj* only with the meaning 'tell to do something, order'.

The verbs that can take both the infinitive and the subordinator are 'want', 'must', 'love', 'let', and 'agree'. In the case of *waržən* 'love' the distribution of these two complementation markers is based on the semantic opposition of 'event' vs. 'proposition', as illustrated by (41). At the same time, with the predicate *ražə wɜvən* 'agree' their distribution is based on coreferentiality: the infinitive is used if its subject is coreferential with the subject of the CTP, otherwise the subordinator *sɜmɜj* is used:

(42) a. *žawər š-ražə iš žɜlinɜ-mɜ a-sɜw-**ən***
Zaur PRVB-agree EXST Zalina-ALL PRVB-go-**INF**
'Zaur agreed to go to Zalina.'
b. **žawər š-ražə iš,* [***sɜmɜj*** *žɜlinɜ-mɜ a-sɜw-a*]
Zaur PRVB-agree EXST **PURP** Zalina-ALL PRVB-go-SBJV.3SG
'Zaur agreed to go to Zalina.'
c. *fəd š-ražə iš,* [***sɜmɜj*** *jɜ= čəžg jɜ= mad-ə*
father PRVB-agree EXST **PURP** his girl her mother-GEN
fsəmɜr-mɜ ba-žžaj-a]
brother-ALL PRVB-stay-SBJV.3SG
'Father permitted his daughter to stay with her uncle.' (lit. 'agreed that his daughter stays')

As can be seen from (42b), the subordinator is unacceptable in case of coreferentiality.

With the verbs *fɜndən* 'want', *qɜwən* 'must' and *bar dɜttən* 'let' the situation is different: the infinitive can only be used if the subject (or experiencer) of the matrix clause is coreferential with the subject (with 'want' and 'must') or object (with 'let') of the complement. The subordinator can be used without any restriction in relation to coreferentiality; cf. (43a) and (43b).

(43) a. *lɜppu-jə* *təng* *fɜndəd-i* [*čəžg-imɜ* *a-qaž-**ən***]
boy-GEN very want-PST.3SG girl-COM PRVB-play-**INF**
'The boy wanted very much to play with the girl.' (Vydrin 2011: 297)
b. *lɜppu-jə* *təng* *fɜndəd-i,* [***sɜmɜj*** *čəžg-imɜ* *a-qažəd-aid*]
boy-GEN very want-PST.3SG **PURP** girl-COM PRVB-play-CNTRF.3SG
'The boy wanted very much to play with the girl.' (Vydrin 2011: 297)

The factors responsible for the relative distribution of the infinitive and the subordinator with these CTPs remain unclear. Vydrin (2011) shows that the choice of the construction with 'want' is not based on the intensity of the desire, knowledge about the realization of the wish, or the truth value of the matrix clause. Based on the use of the subordinator *sɜmɜj* with other CTPs, I suggest that the subordinator *sɜmɜj* here may only introduce a proposition, while the infinitive can encode both events and propositions. However, this differentiation is rather subtle and hard to verify in this context.

The subordinator *kʷəd* ('how')[6] in its second meaning 'in order that' marks strict orders, with three CTPs: 'want', 'must' and 'say':

(44) *armimaz* *ɜmɜ* *ardaʁuj-ɜn* *žɜʁ,* [*a-rdɜm* ***kʷəd***
Armimaz and Ardaguj-DAT say[IMP.2SG] DEM.PROX-DIR **in.order.that**
*ra-sɜw-**oj***], *təng* *žɜrdiag* *qʷəddag-ɜn* *=mɜ* *qɜw-ənʃ...*
PRVB-go-**SBJV**.3PL very important business-DAT me.GEN need-PRS.3PL
'Tell Armimaz and Ardaguj to come here by all means: I need them very much for a certain important business.'
(ONC)

(45) *mɜn* *fɜnd-ə,* [*dɜš* *šaχat-əl* ***kʷəd*** *ɜrba-sɜw-**aj***]
I.GEN want-PRS.3SG ten hour-SPESPL **in.order.that** PRVB-go-**SBJV**.2SG
'I require that you come at 10 o'clock.' (an order)

The subordinators described in this section differ from *kɜj* and *kʷə* in that they require non-indicative mood in the complement verb. Most often this is the subjunctive, as in (44) and (45); however, the counterfactual mood occurs if the situation in the complement clause contradicts the speaker's actual knowledge (see Vydrin 2011 for details):

(46) *mɜn* *fɜndəd-i* [***sɜmɜj*** *=mən* *ba-χːʷəš* *kotː-**aiš*** *fɜlvarɜn-mɜ*
I.GEN want-PST.3SG **PURP** me.DAT PRVB-help do-**CNTRF.**2SG exam-ALL
ba-sɜtːɜ *kɜn-ən*]
PRVB-prepare do-INF
'I wanted you to help me to prepare for my exams' (but you didn't, and I failed).

6 In this function it can occur together with the subordinator *sɜmɜj*.

3.3.4 The conjunction *ӕmӕ* 'and'

The conjunction *ӕmӕ* is most often described as a coordinating conjunction with subordinative functions (Abaev 1950: 656; Kulaev 1959: 72–76; Gagkaev 1956: 222). It coordinates NPs, verbs, and clauses and is also used as a subordination marker, occurring in combination with other subordinators, as *ӕmӕ k^{w}ӕd* in (47), with correlative pronouns, as in (48), or alone, as in (49).

(47) *mӕn* *fӕnd-ə,* [***ӕmӕ*** *dӕš* *šaχat-ӕl* ***k^{w}ӕd*** *ӕrba-sӕw-aj*]
I.GEN want-PRS.3SG **and** ten hour-SPESPL **in.order.that** PRVB-go-SBJV.2SG
'I require that you come at 10 o'clock.'

(48) *henər* *ӕž* ***wəj*** *žon-ən,* [***ӕmӕ*** *sӕwa-jӕ* *nik^{w}ədӕmwal*
now I **that.DEM** know-PRS.1SG **and** Sawa-ABL nowhere
a-irvӕž-zən-ӕn]
PRVB-escape-FUT-1SG
'Now I know that I will not escape from Sawa anywhere.'
(Nart sagas)

(49) *ӕnqӕl* *dӕn,* [***ӕmӕ*** *je=mbal-mӕ* *a-səd-iš*]
think be.PRS.1SG **and** 3SG.POSS=friend-ALL PRVB-go-PST.INTR.3SG
(Where is your brother? – He's not here) 'I think he went to his friend.'

The position of the conjunction *ӕmӕ* is not the same as the position of subordinators (Belyaev 2011; Belyaev 2014a). In complementation, as well as in coordination, *ӕmӕ* can only occur between the two clauses and never at the beginning of the whole sentence or preverbally. This is a strict rule, which is never violated. The order of the clauses is also strictly fixed: unlike all other subordinators, *ӕmӕ* requires in complementation that the matrix clause be preposed to the complement, as in (47)–(49).

In complementation, the conjunction *ӕmӕ* is used in the following contexts. a) It is used to encode propositions with the verbs of speech *žӕʁən* 'say' and *žӕrdӕ ӕvӕrən* 'promise', the mental CTPs *wərnən* 'believe', *aftӕ kӕšən* 'think', *ӕnqӕlən* 'think' and *ӕnqӕl wӕvən* 'think', and the emotive CTPs *žӕrdӕ darən* 'hope', *tӕršən* 'fear', *žӕrdӕmӕ sӕwən* 'like' and *waržən* 'love'.

Without the correlate, these constructions introduce propositions. (49) and (50) are examples of this function: the complement clause contains new information and belongs to the assertion.

(50) *kʷəd* *=dɜm* *kɜš-ə,* *fɜlvarɜn-t-ɜ* *rat-zɜn?* *ɜvi*
how thee.ALL seem-PRS.3SG exam-PL-NOM.PL give-FUT[3SG] or
=dɜm *aftɜ* *kɜš-ə,* [***ɜmɜ*** *dəww3* *ra-jš-zɜn*]*?*
thee.ALL so seem-PRS.3SG **and** two PRVB-take-FUT[3SG]
'What do you think, will he pass the exam? Or you think he's going to fail' (lit. take a two)?

With the correlate, the conjunction *ɜmɜ* introduces topical or previously mentioned propositions, as shown in Section 3.3.8.6.

b) It is used to encode propositions with the verbs *fɜndən* 'want', *wazən* 'let', and *žɜʁən* 'say' in the meaning 'tell to do something.'; in this case the complement verb is in the subjunctive mood:

(51) *mɜn* *fɜnd-ə,* [***ɜmɜ*** *də* *žar-aj*]
I.GEN want-PRS.3SG **and** thou sing-SBJV.2SG
'I want you to sing.'

The conjunction *ɜmɜ* is more general than *kʷəd* (recall [45]) with these verbs: it can introduce both orders and wishes, whereas *kʷəd* marks strict orders.

c) Another use concerns the combination of the conjunction *ɜmɜ* with complement subordinators, such as *kɜj* in (52) and *kʷəd* in (53).

(52) a. *ažɜmɜt* *š-ražə* *iš,* [*rašt* ***kɜj*** *nɜ* *wəd*], *u-wəl*
Azamat PRVB-agree EXST right **COMP** NEG be[PST.3SG] that.DEM-SPESPL
'Azamat agreed that he was wrong.'
b. *ažɜmɜt* *š-ražə* *iš,* [***ɜmɜ*** *rašt* ***kɜj*** *nɜ* *wəd*],
Azamat PRVB-agree EXST **and** right **COMP** NEG be[PST.3SG]
u-wəl
that.DEM-SPESPL
'Azamat agreed that he was wrong.'

(53) a. *mɜn* *fɜnd-ə,* [*dəš* *šaχat-əl* ***kʷəd*** *ɜrba-sɜw-aj*]
I.GEN want-PRS.3SG ten hour-SPESPL **in.order.that** PRVB-go-SBJV.2SG
'I require that you come at 10 o'clock.'
b. *mɜn* *fɜnd-ə,* [***ɜmɜ*** *dəš* *šaχat-əl* ***kʷəd*** *ɜrba-sɜw-aj*]= (47)
I.GEN want-PRS.3SG **and** ten hour-SPESPL **in.order.that** PRVB-go-SBJV.2SG
'I require that you come at 10 o'clock.'

The combination of the conjunction with the complement subordinators in (52b) and (53b) is used to focalize the subordinate clause (see Belyaev 2014b). In this function, it occurs not only in complementation, but in all types of subordination.

3.3.5 The subordinator *səma* 'as if'

The subordinator *səma* 'as if' is used to encode irrealis propositions. The speaker uses this subordinator if s/he is convinced that the proposition is false for certain or with a high degree of probability. Consider the following pairs of examples:

(54) a. *asʲa* *aftɜ* *ɜnqɜlt:-a,* [***səma*** *žawər* *ɜgaš* *u*]
Asja so think-PST.3SG **as.if** Zaur alive be.PRS.3SG
'Asja thought Zaur was alive' (but we know that he's not).

b. *asʲa* *aftɜ* *ɜnqɜlt:-a,* [*žawər* *ɜgaš* ***kɜj*** *u*]
Asja so think-PST.3SG Zaur alive **COMP** be.PRS.3SG
'Asja thought Zaur was alive.' (We don't know if this is the case or not.)

(55) a. *ɜnqɜl* *dɜn,* [***səma*** *sɜw-inag* *u*]
think be.PRS.1SG **as.if** go-PTCP.FUT be.PRS.3SG
'I think he might come' (but I'm not at all sure).
b. *ɜnqɜl* *dɜn,* [*sɜw-inag* ***kɜj*** *u*]
think be.PRS.1SG go-PTCP.FUT **COMP** be.PRS.3SG
'I believe he is going to come' (with more confidence).

In (54a) the speaker is absolutely sure that the situation in the complement is not true. This interpretation does not arise with the subordinator *kɜj* in (54b): here the speaker is not making any prediction about the truth of the situation in question. The pair in (55a) and (55b) differ with regard to the epistemic value of the complement: there is more certainty in the (55b) example, where *kɜj* is used, and less certainty in (55a) with *səma*.

The interpretation of the complement clause as false or doubtful depends on the pragmatic context. This subordinator can be characterized as bearing irrealis propositional value; it is unacceptable with factive verbs, such as 'know':

(56) **ɜž žonən səma*...
'I know as if ...'

3.3.6 The subordinators *kɜd* 'if, when' and *salənmɜ* 'until, as long as'

The subordinator *kɜd* is used in temporal and conditional adverbial clauses[7] (cf. Vydrin 2009). In sentential complement constructions, this subordinator can only be used with the CTP 'wait':

7 The temporal/conditional *kɜd* (including *kɜd* with the verb 'wait') is differentiated from *kɜd* in indirect questions (*I do not know when you were born*), which has a strict preverbal position

(57) *šəvɜllon* *ɜnqɜlmɜ* *kašt-i,* [*də* *=jən* ***kɜd*** *ba-χɜr-ən* *kɜn-zən-ɜ*]
child wait look-PST.3SG thou he.DAT **if** PRVB-eat-INF do-FUT-2SG
'The child waited for you to feed him.'

The use of the subordinator *salənmɜ* 'until, as long' in complementation is also restricted to the CTP 'wait':

(58) *maχ* *ɜnqɜlmɜ* *kɜš-ɜm,* [***salənmɜ*** *waẑẑə-t-ɜ* *ɜrba-sɜw-oj*]
we (wait) look-PRS.1PL **until** guest-PL-NOM.PL PRVB-go-SBJV.3PL
'We are waiting for the guests to come.'

That the verb 'wait' and its synonyms require a special complementation construction is a widespread phenomenon in the languages of the world; examples include the subordinator *poka* 'until' in Russian, and the specialized converbs in Qunqi and Xuduc Dargwa (Nakh-Dagestanian; cf. Serdobolskaya 2009).

There is no apparent semantic difference between *kɜd* and *salənmɜ* in complementation. The complement clause of the verb 'wait' can be headed by the verb in the indicative or optative/counterfactual mood; cf. (59) and (60). The latter is chosen if the situation has counterfactual meaning:

(59) *ɜž* *ɜnqɜlmɜ* *kɜš-ən,* [*urok* ***kɜd*** *fɜ-wə-**zɜn***], ***wə-mɜ***
I (wait) look-PRS.1SG lesson **if** PRVB-finish-**FUT[3SG]** **that-ALL**
'I'm waiting for the lesson to end.'

(60) *ɜž* *ɜnqɜlmɜ* *kɜš-ən,* [*Zaur* ***kɜd*** *ɜrba-sɜw-**id***]
I (wait) look-PRS.1SG Zaur **when** PRVB-go-**OPT.3SG**
'I'm waiting, in case Zaur comes. (It was arranged that he would not.)'

For the choice of non-indicative mood (optative or counterfactual) consult Vydrin (2011).

3.3.7 Parataxis

The paratactic construction contains two clauses, matrix and complement, without any overt marker of subordination[8]. The two clauses can occur in either order, cf.:

(like the interrogative *kɜd* as *When were you born?*). By contrast, the temporal/conditional *kɜd* is a 'floating' subordinator.

8 It has not been investigated in detail whether these constructions can be described in terms of subordination; the "root clause" properties (Green 1976) of both clauses remain to be checked.

(61) a. *зž зnqзlt:-on,* [*žawər зgaš u*]
I think-PST.1SG Zaur alive be.PRS3SG
'I thought that Zaur was alive.'
b. [*žawər зgaš u*], *зž зnqзlt:-on*
Zaur alive be.PRS.3SG I think-PST.1SG
'I thought that Zaur was alive.'

The paratactic construction introduces propositions (62) or irrealis complements (63) with non-factive verbs, mental verbs of opinion, speech verbs, emotive verbs and the verb 'intend'. See (62), where the truth of the complement clause is asserted by the speaker, and (63), where it is strongly doubted.

(62) *aχʷərgзnзg žaχt-a,* [*rənčən u*]
teacher say-PST.3SG sick be.PRS.3SG
'(Where is Zaur?) – 'The teacher said he was sick.'

(63) [*žawər зgaš u*], ***wəj*** *зnqзl nз dзn*
Zaur alive be.PRS.3SG **that.DEM** (think) NEG be.PRS.1SG
'I'm not sure that Zaur is alive.'

This construction can take a correlative pronoun as in (63); see also Section 3.3.8.2.

3.3.8 Correlative pronouns/adverbs in complementation

3.3.8.1 Position of the correlative pronouns/adverbs

Ossetic subordinators can occur with correlative pronouns/adverbs in the matrix clause, or without them:

(64) *зž žon-ən,* [*žawər čəžg* ***kзj*** *зr-χašt-a*], ***(wəj)***
I know-PRS.1SG Zaur girl **COMP** PRVB-take-PST.3SG **that.DEM**
'I know that Zaur has married.'

Correlative pronouns/adverbs are traditionally analyzed as components of complex subordinators; e.g. *kзj ... wəj* in (64), *kʷəd... wəj* etc. (Abaev 1950: 718–719; Kulaev 1959; Gagkaev 1956: 222–224). However, as shown in (Belyaev & Serdobolskaya, in project; cf. also Bagaev 1982), it is more appropriate to analyze the subordinators and the correlative pronouns separately: the pairs are not fixed; various correlates can be used with one and the same subordinator, as shown in Section 3.3.2, and the correlate does not even have to be a pronoun: a noun phrase with the demonstrative pronoun/adverb can also serve as a correlate. When two or more subordinate clauses co-occur, the correlate takes the plural form. Thus, the

correlates and subordinators do not form fixed pairs that belong to the lexicon (unlike English *if* ... *then* ... and similar cases).

The correlate must be adjacent to the complement clause. The correlative pronoun *wəj* (3rd person singular) takes the case marker required by the argument structure of the CTP, which is nominative/genitive in (64) and superlative in (65).

(65) [*mɜ=* *mad* *raʒə* ***kɜj*** *šə-št-ə*], ***u-wəl*** *š-aχʷər*
my mother early **COMP** PRVB-stand-PRS.3SG **that-SPESPL** PRVB-study
dɜn
be.PRS.1SG
'I'm used to mother's getting up early.'

As exemplified in Section 3.1, all subordination types in Ossetic make use of constructions with correlative pronouns/adverbs in the matrix clause and subordinators in the dependent clause. However, the possibility of omitting the pronoun, demonstrated by (64), is only attested in complementation and the purpose construction (Belyaev 2011).

Abaev (1950) proposes the following rule for the omission of correlative pronouns in complementation: the pronoun is obligatory if the matrix clause is postposed, and optional if the matrix clause is preposed (Abaev 1950: 719), cf. (64) and (66):

(66) a. [*žawər* *ɜgaš* ***kɜj*** *u*], ***wəj*** *ɜž* *žon-ən*
Zaur alive **COMP** be.PRS.3SG **that.DEM** I know-PRS.1SG
'I know that Zaur is alive.'
b. *[*žawər* *ɜgaš* ***kɜj*** *u*], *ɜž* *žon-ən*
Zaur alive **COMP** be.PRS.3SG I know-PRS.1SG
Intended: 'I know that Zaur is alive.'

Hence, there are three possible constructions with respect to the order of the matrix and the complement clause:

(A) MatrCl DepCl *wəj*
(B) MatrCl DepCl
(C) DepCl *wəj* MatrCl

This rule is strict and it works for all complement subordinators, with the exception of *ɜmɜ*; see Section 3.3.4 above.

It remains unclear what triggers the omission of correlative pronouns if the matrix clause is preposed, i.e. what is the semantic difference between the variants illustrated in (64). It must be admitted that the use of the pronoun does not obey strict grammatical rules, but merely demonstrates strong tendencies.

3.3.8.2 The subordinators *kɜj* and *səma*, and the paratactic construction

With the subordinator *kɜj* the correlative pronoun distinguishes between facts and propositions, e.g. between complement clauses belonging to the presupposition and to the assertion:

(67) *ɜž* *žon-ən,* [*žawər* *ɜgaš* ***kɜj*** *u*], ***wəj***
I know-PRS.1SG Zaur alive **COMP** be.PRS.3SG **that.DEM**
'I know that Zaur is alive.'

(68) *žawər* *kɜm* *iš? –* *aχʷərgɜnɜg* *žaχt-a,* [*rənčən* ***kɜj*** *u*]
Zaur where EXST teacher say-PST.3SG ill **COMP** be.PRS.3SG
'Where is Zaur? – The teacher said that he is ill.'

The verb 'know' in (67) introduces a fact, and the complement takes the correlative pronoun, while in (68), where the truth of the complement is not presupposed, the pronoun does not occur. Paratactic complements that take the correlative pronoun also introduce facts; see (69). It must be admitted, however, that such examples are rarer than examples with *kɜj*.

(69) *asʲa* *ɜnqɜl* *nɜ* *wəd-i,* [*žawər* *ɜgaš* *u*], ***wəj***
Asja think NEG be-PST.3SG Zaur alive be.PRS.3SG **that.DEM**
'Asja didn't even think that Zaur could be alive' (but we know that he is).

Let us consider further evidence for this semantic distinction. First, the correlative pronoun is most often present with factive verbs such as 'know'. With non-factive verbs the correlative pronoun is most often absent in contexts where the dependent clause presents new information, cf. (68) and (70), or in irrealis contexts as in (71).

(70) *radio-jɜ* *ra-zərt:-oj,* [*rajšom* *wažal* ***kɜj*** *wə-zɜn*]
radio-ABL PRVB-tell-PST.3PL tomorrow frost **COMP** be-FUT[3SG]
'They said on the radio that it will be freezing tomorrow.'

(71) *ɜž* *nɜ* *žaχt-on,* [*ražə* *dɜn*], *ɜp:ənd ɜr* *nisə* *š-zərt:-on*
I NEG say-PST.1SG agreeing be.PRS.1SG at.all nothing PRVB-speak-PST.1SG
(Why do you think I agreed?) 'I didn't say I agreed, I remained silent.'

The same CTP 'say' takes the correlative pronoun if the truth of the complement is presupposed:

(72) *dɜ=* *mad-ɜn* *sɜ-wəl* *nɜ* *žaχt-aj,* [*dəwwɜ* ***kɜj***
thy mother-DAT what-SPESPL NEG say-PST.2SG two **COMP**
ra-jšt-aj], ***wəj?***
PRVB-get-PST.2SG **that.DEM**
'Why didn't you tell mother that you got a bad mark?'

Second, the truth of the complement clause cannot be denied in the following context by the same speaker, cf.:

(73) **də žon-əš, [žawər čəžg **kɜj** ɜr-χašt-a], **wəj**, fɜlɜ*
you know-PRS.2SG Zaur girl **COMP** PRVB-take-PST.3SG **that.DEM** but
wəj ɜsɜg nɜ-w
that.DEM true NEG-be.PRS.3SG
Intended meaning: 'You know that Zaur has married, but this is not true.'

This sentence is possible in the context of non-factive verbs and without the correlative pronoun.

Third, in performative and quasi-performative contexts the complement clause without the correlative pronoun must be used:

(74) *quš-ən =dən kɜn-ən, [nər-ɜj fɜštɜ-mɜ am nal*
listen-INF thee.DAT do-PRS.1SG now-ABL later-ALL here no.more
kuš-əš] (***wəj**)
work-PRS.2SG **that.DEM**
'I declare that you're fired.' (pronounced by an authorized person)

Fourth, complement clauses with factive verbs such as 'know' do not take correlative pronouns in the protasis of counterfactual conditionals:

(75) *k^{w}ə =jɜ žon-in, [amondžən **kɜj** wə-zən-ɜ], wɜd =wɜ*
if it.GEN know-OPT.1SG happy **COMP** be-FUT-2SG then you.GEN
nɜ qəgdar-in
NEG object-OPT.1SG
'If I knew that you were happy, I wouldn't hinder [this marriage]' (but I do hinder it as I don't believe you are).'

By contrast, in (76) the presupposition is kept (as can be seen from the context) and the correlate must be used.

(76) *ba-sin kotː-ain, [zawər **kɜj** ɜrba-səd-i], **u-wəl***
PRVB-joy do-CNTRF.1SG Zaur **COMP** PRVB-go-PST.3SG **that-SPESPL**
(Zaur came yesterday, it is a pity you didn't see him.) '– Oh yes, I would have been glad he had come!' (I am not glad now, since I haven't seen him.)

Fifth, the subordinator *səma*, which introduces propositions but cannot introduce facts, does not take correlates in complementation:

(77) a. *as^{j}a aftɜ ɜnqɜl-ə, [**səma** žawər ɜgaš u], (??**wəj**)*
Asya so think-PRS.3SG **as.if** Zaur alive be.PRS.3SG **that.DEM**
'Asya believes that Zaur is alive.'
b. *[**səma** žawər ɜgaš u], (??**wəj**) as^{j}a aftɜ ɜnqɜl-ə*
as.if Zaur alive be.PRS.3SG **that.DEM** Asja so believe-PRS.3SG
'Asya believes that Zaur is alive.'

As stated in Section 3.3.5, this subordinator is used to introduce non-factive complements: hence the incompatibility with the correlative pronoun[9].

This evidence sustains the claim that the correlative pronoun is used in cases where the complement clause belongs to the presupposition. This claim is in line with the word order rule formulated by Abaev; cf. (A)–(C) in Section 3.3.8.1 above. The dependent clause is often preposed to the matrix if it constitutes the topic of the sentence. In this position it obligatorily takes the correlative pronoun. That means that topical clauses take the correlative pronoun. Topical complement clauses are often encoded in the same way as presupposed complements (see the discussion of the Adyghe data in Serdobolskaya, this volume).

Just like factive clauses, irrealis clauses in topical position take the correlative pronouns:

(78) *зӕ* *nз* *fe-q^{w}əšt-on*, [*χəl* ***kɜj*** *kot:-oj*], ***wəj***, *ɜmɜ*
I NEG PRVB-hear-PST.1SG quarrel **COMP** do-PST.3PL **that.DEM** and
mɜ *nз* *wərn-ə*
me.GEN NEG believe-PRS.3SG
(Zaur quarreled with his wife!) '– I haven't heard that they quarreled, and I don't believe it.'

The context shows that the complement clause in (78) is interpreted as false. Hence it represents an example of an irrealis proposition and not a fact. However, the correlative pronoun is used, because the complement clause is topical.

The same distribution is observed with the paratactic construction, cf. Section 3.3.7: if the complement clause is postposed to the matrix, the correlative pronoun signals its presupposed (recall [69] above) or topical status. What concerns preposed complement clauses, the correlative pronoun is not obligatory in parataxis, unlike in the construction with conjunctions. If the correlate is present in the sequence 'complement + matrix clause', it most often signals that the complement is topical:

9 In adverbial clauses, this subordinator can take correlates, e.g. the demonstrative pronoun in the equative case:

(i) *...žaχt-a* *farnɜg*, [***səma*** *wə-mɜ* *nз* *zərt-a*, *zɜmbə-mɜ*,
say-PST.3SG Farnag **as.if** that.DEM-ALL NEG speak-PST.3SG Dzamby-ALL
səma *ɜnɜwi* *žaχt-a*], ***wəj-aw***.
as.if for.no.particular.reason say-PST.3SG **that.DEM-EQU**
(Lumps of coal are only good in the fireplace), '– Farnag said, as if he were not speaking to Zamba, as if he were saying it for no particular reason'
(lit. as if just said, this way).
(Nart sagas)

(79) [*žawər ʒgaš u*], ***wəj*** *ʒnqʒl nʒ dʒn*
Zaur alive be.PRS.3SG **that.DEM** think NEG be.PRS.1SG
(Is Zaur alive?) '– I don't not think that Zaur is alive.'

With the remaining subordinators, the factor of presupposition/assertion is not relevant. The semantic opposition of the complements with and without the correlate is based on the opposition of 'old vs. new' or 'expected vs. unexpected' information. This opposition is formed directly by the pragmatic context, and hence, the semantic difference between the sentences with and without the correlate is often subtle and not easy to deduce by elicitation or with the help of corpus examples. Similar generalizations have been made on the use of the correlate *es* in German complements (see Dalmas 2013): the optionality of this correlate is dealt with in terms of topicality and/or mentionedness of the situation in the complement clause.

As the correlate is obligatory with all subordinators if the complement is preposed, in what follows I only consider examples with postposed complements.

3.3.8.3 The subordinators *k^{w}ə* and *k^{w}əd*

The correlate is used with the subordinators *k^{w}ə* and *k^{w}əd* to introduce old information, present in the preceding discourse or pragmatic context. Compare (80), where the complement has been introduced earlier in the discourse and accordingly the correlative pronoun is present, and (81), where the speaker gives an opinion that has not been discussed before.

(80) *zawər ʒmʒ jʒ= uš fʒ-χəl štə! – ʒž =ʒj nʒ fe-q^{w}əšt-on,*
Zaur and his wife PRVB-quarrel be.PRS.3PL I it.GEN NEG PRVB-hear-PST.1SG
[***k^{w}əd*** *χəl kotː-oj*], ***wəj***, *ʒmʒ =mʒ nʒ wərn-ə*
how quarrel do-PST.3PL **that.DEM** and me.GEN NEG believe-PRS.3SG
'Zaur has quarrelled with his wife! – I haven't heard them quarrelling, and I don't believe it.'

(81) *wən-əš,* [***k^{w}əd*** *=nʒ šaj-ənsˈ*]*!*
see-PRS.2SG **how** us.GEN deceive-PRS.3PL
(Now you have learnt how the men love!) 'You see how they deceive us!'
(ONC)

The same tendency is observed with the subordinator *k^{w}ə*: see (82), where the complement has been mentioned previously, and the correlative pronoun is used, and (83), where the complement belongs to the new information: the speaker introduces his/her request, and his/her fear that this will not be accepted.

(82) *žon-ə* *=jɜ,* *bɜrgɜ,* *jɜ=* *χi* *kʷə* *r-waz-id,* *wɜd* *=ən*
know-PRS.3SG it.GEN certainly his self if PRVB-let-OPT.3SG then he.DAT
kɜj *fe-nson-dɜr* *w-aid,* *wəj,* *fɜlɜ* *=jɜ* *aftɜ* *=dɜr* *nɜ*
COMP PRVB-easy-CMP be-OPT.3SG that.DEM but he.GEN so ADD NEG
fɜnd-ə, *wəm-ɜj* *=dɜr* *tɜrš-ə,* [***kʷə*** *=jən* *fe-nson-dɜr*
want-PRS.3SG that-ABL ADD fear-PRS.3SG **when** he.DAT PRVB-easy-CMP
w-a], ***wəm-ɜj***
be-SBJV.3SG **that-ABL**
'Certainly he knows that it would be easier for him if he lay down (for a rest; lit. if he let himself), but he doesn't want that, he's even afraid of it being easier for him.'
(ONC)

(83) *fɜlɜ* *=mɜm* *iw* *kurdiat* *iš,* *ɜmɜ* *tɜrš-ən,* [***kʷə*** *nɜ*
but me.ALL one request EXST and fear-PRS.1SG **when** NEG
š-ražə *w-aj*]
PRVB-agreeing be-OPT.2SG
'I have a request, and I'm afraid that you won't consent.'
(ONC)

3.3.8.4 The subordinators *kɜd* and *salənmɜ*

The subordinators *kɜd* and *salənmɜ* are used only with the verb 'wait' (and its synonyms). The rationale for the use of correlates with these subordinators is more transparent than with *kʷə* and *kʷəd*. The correlate is used if the complement clause denotes an event that surely occurs or has occurred ('wait until the end of summer, the end of the lesson etc.'):

(84) *ɜnqɜlmɜ* *kašt-əštɜm,* [*war-ən* ***kɜd*** *ba-nsaj-zɜn*], ***wə-mɜ***
(wait) look-PST.1PL rain-INF **if** PRVB-stop-FUT[3SG] **that-ALL**
(The fire has built up and our wet clothes have dried.) 'We were waiting until the rain stops.'
(ONC)

(85) *dɜw* *kʷə* *nɜ* *fet-ain,* *wɜd* *=ma* *=mɜ* *birɜ*
thee.GEN if NEG see.PFV-CNTRF.1SG then also me.GEN long
ba-qʷəd-aid *ɜnqɜlmɜ* *kɜš-ən,* [***salənmɜ*** *bazar* *ba-jgom*
PRVB-must-CNTRF.3SG (wait) look-INF **until** bazaar PRVB-open
wəd-aid], ***wɜd-mɜ***
be-CNTRF.3SG **then-ALL**
'If I hadn't seen you, I would have had to wait a lot until the bazaar opens.'
(ONC)

In (84) and (85) the correlate is used, since it is common knowledge that the rain will eventually stop, and the bazaar opens every morning. By contrast, the complement clause occurs without the correlate if the actualization of the event denoted by the complement is not certain (or even impossible) for pragmatic reasons:

(86) *kʷəz* *kʾəšər-ə* *šər-t-ə* *ba-gəp:* *kot:-a...* *ənqəlmə* *kašt-i,*
dog threshold-GEN top-PL-IN PRVB-jump do-PST.3SG (wait) look-PST.3SG
[***kəd*** *=ən* *ištə* *χərinag* *ra-p:ar-ik:oj*]
if it.DAT something food PRVB-throw-OPT.3PL
'The dog jumped over the threshold and inside the house, waiting for somebody to throw it some food.' (No, nobody did, those two people were too busy.)
(ONC)

(87) *ənqəlmə* *kəš,* [***salənmə*** *=dən* *wərəš-ə* *pac:aχ* *jə=* *bandon*
(wait) look[IMP.2SG] **until** thee.DAT Russia-GEN emperor his throne
əfštaw *a-vər-a*]
on.credit PRVB-give-SBJV.3SG
'Wait for the Russian emperor to lend you his throne.'
(ONC)

Example (87) is especially telling, because the complement denotes an event that is pragmatically impossible.

This semantic opposition influences also the choice of the mood of the verb in the complement clause, cf. Section 3.3.6. Therefore, it is not surprising that there is a correlation between the presence of the correlate and the choice of mood: the indicative mood is most often observed if the correlate is present, while non-indicative moods are mostly attested if the correlate is absent.

3.3.8.5 The subordinator *səmə*j

Usually this subordinator does not take a correlate: out of 63 examples of *səmə*j in complement clauses (with a postposed complement clause), arbitrarily taken from the corpus, only four contain the correlate. In all those examples the complement denotes a situation that is somehow discussed in the previous context[10], e.g.:

(88) *aχχošəg-t-əj* *=ma* *iw* *wəj* *u,* *əmə* *səvəllə-t:-ə*
reason-PL-ABL PRT one that.DEM be.PRS.3SG and child-PL-GEN
nəjjaržə-t-əj *=dər* *kəj-dər-tə* *nə* *fə-fənd-ə,* [***səmə*j**
parent-PL-ABL ADD who.GEN-INDF-PL.OBL NEG PRVB-want-PRS.3SG **PURP**
jə= *šəvəllon* *iron-aw* *aχʷər* *kən-a*], ***wəj***
his child Ossetian-EQU learn do-SBJV.3SG **that.DEM**
'There is one more reason: some of the children's parents don't want their children to learn the Ossetic language.'
(ONC)

10 I thank Oleg Belyaev for this calculation and the example.

The sentence in (88) comes from an article ("On Ossetic – from the point of view of the law"; magazine *Max dug*, 2006: 5), which discusses the problems of speaking Ossetic in nurseries and kindergartens, and the situations of "learning Ossetic" and "speaking Ossetic" are thus mentioned repeatedly throughout the article.

The complement in (89) is an answer to a question and, hence, presents new information. Therefore, the correlate is not used:

(89) *wɜdɜ* *=dɜ* *kʷəd* *fɜnd-ə? –* *mɜn* *fɜnd-ə,* [***sɜmɜj*** *tuʁan-ə*
then thee.GEN how want-PRS.3SG I.GEN want-PRS.3SG **PURP** Turan-GEN
ma-či *šš-ar-a*]*!*
NEG-who PRVB-find-SBJV.3SG
'What do you want then? – I want no one to find Turan.'
(ONC)

3.3.8.6 The conjunction *ɜmɜ*

The conjunction *ɜmɜ* does not show the same syntactic properties as other complement subordinators. First, the complement with the conjunction *ɜmɜ* cannot appear preposed to the matrix clause (cf. construction C in Section 3.3.8.1). Second, the correlate can either be placed inside the matrix clause (this construction does not occur with other subordinators) or postposed to the complement clause (cf. construction A in Section 3.3.8.1):

(90) *henər* *ɜž* ***wəj*** *žon-ən,* [***ɜmɜ*** *sɜwa-jɜ* *nikʷədɜm-wal*
now I **that.DEM** know-PRS.1SG **and** Sawa-ABL nowhere-more
airvɜž-zən-ɜn] =(48)
escape-FUT-1SG
'Now I know that I will not escape from Sawa anywhere.'
(Nart sagas)

(91) *sɜwən* *nɜ* *žaχt-aj,* [***ɜmɜ*** *dɜ=* *mad-ɜn* *ɜχχʷəš* *kot:-aj*],
why NEG say-PST.2SG **and** thy mother-DAT help do-PST.2SG
wəj*?*
that.DEM
(Teacher scolded me for not having done my homework. – You should have lied.) 'Why didn't you say that you had been helping your mother?'

However, it is noteworthy that the correlate is only rarely observed with *ɜmɜ*, which is not surprising given that this conjunction serves to introduce propositions.

In spite of the syntactic differences between the construction with *ɜmɜ* and those in which the other subordinators occur, the semantic difference associated with the presence vs. absence of the correlate is the same with *ɜmɜ* as was observed above for the constructions with other subordinators. The correlate is

present if the complement is topical or refers to old/expected information (90). For example, the correlate is rejected by native speakers if the complement constitutes the focus:

(92) *ɜnqɜl* *dɜn,* [***ɜmɜ*** *je=* *mbal-mɜ* *a-səd-iš*] (****wəj***)
think be.PRS.1SG **and** his friend-ALL PRVB-go-PST.3SG **that.DEM**
'{Where is your brother?} – I think he went to his friend.'

3.3.9 The citation particles *dam* and *žɜʁgɜ*

Žɜʁgɜ is a participle-converb of the speech verb *žɜʁən*. It is used as a complementizer, mostly to indicate propositions, with various classes of CTPs, cf.:

(93) [*χɜzar* *nal* *wɜj* *kɜn-ən,* ***žɜʁ-gɜ***], *ɜž* *žaχt-on*
house no.more purchase do-PRS.1SG **say-CVBP** I say-PST.1SG
'I said that I wouldn't sell the house.'
(TEXT)

(94) *ɜž* *fe-q^{w}əšt-on,* [*mɜšk^{w}ə-jə* *mit* *war-ə,* ***žɜʁ-gɜ***]
I PRVB-hear-PST.1SG Moscow-IN snow fall-PRS.3SG **say-CVBP**
'I've heard that it's snowing in Moscow.'

Žɜʁgɜ is grammaticalized as a complementizer. First, it can appear alongside the verb *žɜʁən* in one and the same sentence, see (93), without giving rise to a tautology. Second, it can be used with CTPs that do not denote speech acts:

(95) *wəj=jɜ* *fe-nqɜlt:-a* [*arš* *u,* ***žɜʁ-gɜ***], *ɜmɜ* *jɜ=* *ražmɜ*
he=it.GEN PRVB-think-PST.3SG bear be.PRS.3SG **say-CVBP** and his near
ra-səd-i.
PRVB-go-PST.3SG
'He thought it was a bear and went before it.'
(TEXT)

The complementizer *žɜʁgɜ* can be used with the imperative in the complement clause to denote orders and requests:

(96) *mɜ=* *fəd* *=mən* *žaχt-a,* [*don* *ra-χɜš,* ***žɜʁ-gɜ***]
my father me.DAT say-PST.3SG water PRVB-carry[IMP.2SG] **say-CVBP**
'My father asked me to bring some water.'

For further details about the distribution of *žɜʁgɜ* consult Vydrin (forthcoming).

The particle *dam* is mostly used with verbs of speech:

(97) *hemɜ* *=iw* *=šən,* [*ɜž* ***=dam*** *=wɜm* *=šɜ* *ba-tɜr-zən-ɜn*].
then ITER they.DAT I **CIT** you.ALL they.GEN PRVB-drive-FUT-1SG
'Then he said: "I'll drive them [the Caucasian goats] to you".'
(TEXT)

Dam can be repeated several times in a single clause (98) or in combination with *žɜʁgɜ* (99):

(98) *ɜmɜ* *rajɜ* *aftɜ* *žɜʁ-ə,* [*wɜd* ***=dam*** *am* ***=dam*** *propiskɜ*
and Raya so say-PRS.3SG then **CIT** here **CIT** registration
š-kɜ].
PRVB-do.IMP.2SG
'And Raya says to me: "Then register here". '
(TEXT)

(99) [*naššən-t-ɜ* *χɜr-ɜn* *zəχ* ***=dam*** ***žɜʁ-gɜ***], *ɜvɜsːɜgɜn,* *aχɜm*
pumpkin.pie-PL-NOM.PL eat-ADJ mouth **CIT** **say-CVBP** probably such
štər *zəχ-ɜj* *fɜ-žɜʁ-ənsˑ*
big mouth-ABL PRVB-say-PRS.3PL
'Such a big mouth they probably call "a pumpkin pie eating mouth". (lit. Of such a big mouth they probably say "pumpkin pie eating mouth")'
(ONC)

Both citation particles can occur in non-subordinate clauses to refer to a citation.

3.4 Non-finite complements: morphosyntax and functional distribution

3.4.1 Infinitive

3.4.1.1 Morphosyntactic properties of the infinitive

The Ossetic infinitive shows the morphosyntactic properties of a verbal noun: it can take all the case markers, the nominal number suffix, and possessive clitics available to nouns (Abaev 1950), as in the following examples:

(100) *fɜšš-**ən**-ɜj* *ba-fɜllad-tɜn.*
write-**INF**-ABL PRVB-be.tired-PST.1SG
'I'm tired of writing.'
(Abaev 1950: 614)

(101) *sɜr-**ən**-t-ɜ* *ba-jdətː-oj.*
live-**INF**-PL-NOM.PL PRVB-begin-PST.3PL
'They began to live well.'
(Abaev 1950: 614)

(102) *mɜ=* *žɜrdɜ-mɜ* *sɜw-ə* [*dɜ=* *kaf-**ən***].
my heart-ALL go-PRS.3SG thy dance-**INF**
'I like your dancing.'

The infinitive does not show tense and mood differentiation. It preserves the perfectivity distinction encoded by preverbs; e.g. *kɜnən* vs. *š-kɜnən* 'do' vs. 'have done'. The infinitive cannot take the finite negation particle *nɜ*; the non-finite negation *ma* can be used with some CTPs:

(103) *alinɜ* *žɜrdɜ* *ba-vɜrt:-a* *nana-jɜn* [*jɜ=* *nog* *k'aba* *ma*
Alina heart PRVB-put-PST.3SG mother-DAT her new dress PROH
š-č'izi *kɜn-**ən***].
PRVB-soil do-**INF**
'Alina promised (lit. put the heart) her mother not to soil her new dress.'

The infinitive preserves the active/passive voice distinction; the passive forms of the infinitive can be found in Caucasian texts:

(104) *"Irəšton-ə"* *volejbolist-t-ɜ* *k^{w}əd* *žɜʁ-ənS,* *aftɜmɜj* *šɜ=*
Iriston-GEN volleyball.player-PL-NOM.PL how say-PRS.3PL this.way their
bon [***fɜ-χɜrd*** ***wɜv-ən***] *nɜ* *wəd...*
possibility **PRVB-beat.PTCP.PST** **be-INF** NEG be[PST.3SG]
'"Iriston" volleyball players said that they could not be beaten' (lit. it was not their possibility to be beaten) (because their victory was a birthday present to their coach Felix Khamikoev.)
(http://alaniatv.ru/habaerta/vesti-iryston/?id=6100)

The infinitive cannot have a subject in the nominative; the subject of the infinitive may only be expressed via genitive (possessive) clitics; see (102). The direct object and the other arguments of the infinitive are expressed in the same way as in the corresponding finite clause:

(105) *mɜn* *fɜnd-ə* [*χɜrinag /* **χɜrinaǯ-ə* *kɜn-**ən***].
I.GEN want-PRS.3SG dinner(NOM) dinner-GEN do-**INF**
'I want to cook food.'

In Ossetic, non-animate direct objects occur most often without any overt marker, while animate DOs take genitive marking (Abaev 1950), as do dependent nominals in NPs. Hence, the unacceptability of the genitive in (105) signals that the infinitive marks its DO in the same way as in the corresponding finite clause. It follows that the infinitival clause preserves the verbal argument structure, except from the marking of the subject and negation; morphosyntactically, however, the infinitive shows nominal properties.

3.4.1.2 Semantics of the infinitive

The distribution of the infinitive is very similar to the distribution of infinitives in English or Russian. It is used to encode complements with future or generic reference, controlled or caused by the subject (or experiencer) of the matrix clause:

(106) *čəžg* *ra-jdət:-a* *kɜw-**ən***.
girl PRVB-begin-PST.3SG weep-**INF**
'The girl started crying.'

(107) *jɜ=* *bon* *u* [*rɜšuʁd* *kaf-**ən***].
her possibility be.PRS.3SG beautiful dance-**INF**
'She can dance beautifully.'

Unlike the infinitives of other languages, the Ossetic infinitive does not encode complements of the verb 'finish', 'end'; instead the nominalization is used:

(108) *lɜp:u* [*kaš* *χɜr**d***] *fɜs-i*.
boy porridge eat:**PTCP.PST** PRVB-EXST
'The boy ate up the porridge (lit. finished eating).'

With control predicates, the infinitive is used if the coreferentiality pattern corresponds to the default for the particular CTP involved. This is subject control for 'promise', and object control for 'let'.

Semantically, the infinitive can encode both events (including generic events) and propositions with control verbs. With evaluative predicates it encodes generic events – see (109) with the predicate *χorž*, where the infinitive contrasts with the construction with *kɜj* and with the correlative pronoun:

(109) a. *šɜrdəgon* *χorž* *u* [*χɜχ-t-ə* *težʁo* *kɜn-**ən***].
in.summer good be.PRS.3SG mountain-PL-IN walk do-**INF**
'In summer, it is nice to take a walk in the mountains.'
b. *sə* *χorž* *u,* [*fən* ***kɜj*** *wəd-iš*], ***wəj***!
what good be.PRS.3SG dream **COMP** be-PST.3SG **that.DEM**
'It is so good that it was a dream!'
(ONC)

The sentence in (109a) describes the feelings experienced about the situation of walking, while in (109b) it is the fact of the situation being true that is evaluated as positive. Hence, (109a) presents an eventive context, and the infinitive is used, while (109b) presents a factive context, expressed by means of the subordinator *kɜj* and the correlative pronoun.

Note, however, that the infinitive is not used to mark events associated with verbs of immediate perception.

With the verbs 'let', 'promise' and others the infinitive encodes propositions:

(110) *žawər* *nɜ* *waz-ə* *je=* *fšəmɜr-ə* [*jɜ=* *χɜsɜngarž-mɜ* *ɜvnal-**ən***].
Zaur NEG let-PRS.3SG his brother-GEN his gun-ALL touch-**INF**
'Zaur doesn't let his brother take his gun.'

The infinitive can be used with verbs of speech only if these function as manipulative predicates and imply causation, as in *I told him to go*.

Thus, the distribution of the infinitive is governed not by the semantics of the complement, but by the coreferentiality pattern and by the presence of causation. For the relative distribution of the infinitive and the complement clauses with *sɜmɜj* see Section 3.3.3.

3.4.2 Nominalization

3.4.2.1 Morphosyntactic properties of the nominalization

The nominalization construction is headed by the participle in *-t/-d* (111). It can take all the morphological markers available to nouns (Abaev 1950); cf. the possessive clitic in (111), the nominal plural marker in (112), and the case marker in (113).

(111) [***de**=* *rba-sə**d***] *=mən* *ɜχšəžgon* *u.*
thy PRVB-go:**PTCP.PST** me.DAT joy be.PRS.3SG
'I'm glad that you've come.'

(112) [*jɜ=* *kɜn**d**-**t**ə**t**-**ɜ***] *=mɜ* *nɜ* *qɜw-ən**s**.*
her do:**PTCP.PST-PL-NOM.PL** me.GEN NEG need-PRS.3PL
'I don't need what she has done (the things that she has done).'

The nominalization does not take the markers of verbal morphological categories, such as mood and tense, and it does not show voice distinction. It can, however, take perfective preverbs; cf. (111). Negation (whether expressed by the indicative negation particle *nɜ* or the modal negation particle *ma*) is also impossible in nominalizations[11].

Syntactically, nominalizations behave like nouns. The semantic subject of the nominalization appears in the genitive:

(113) *zž* [*birɜʁ-**ə** /* **birɜʁ* *niw**ə**d-**ɜj***] *tɜrš-ən.*
I wolf-**GEN** wolf(NOM) howl:**PTCP.PST-ABL** fear-PRS.1SG
'I fear the wolf's howl.'

11 In some contexts the corresponding meaning can be expressed by the preposition *ɜnɜ* 'without' (e.g. 'One cannot live **without eating**'). However, since this is not possible in complementation, we do not consider such examples here.

(114) [*mašinɜ-**jə** / *mašinɜ ɜlχɜ**d**-ə fɜštɜ] *avtobus-əl nal sɜw-ɜm.*
car-**GEN** car(NOM) buy:**PTCP.PST**-GEN after bus-SPESPL no.more go-PRS.1PL
'Since buying the car we do not take the bus any more.'

The direct object can only occur in the genitive, unlike in the corresponding independent clauses (where the genitive is mostly used for animate and the nominative for non-animate DOs). It is impossible for nominalizations to take both a subject and an object; only one of these is acceptable in nominalized clauses (unless the subject is expressed by means of a possessive pronoun). Circumstantials can only be encoded by adjectives: for example, temporal adverbs must take the genitive that functions as adjectivizer:

(115) *alinɜ-jə žnon-**ə** / *žnon kaf**t***
Alina-GEN yesterday-**GEN** yesterday dance:**PTCP.PST**
'Alina's dance yesterday', lit. 'yesterday's dance of Alina'

Therefore, nominalizations behave like nominals with regard to their morphosyntactic properties: they manifest nominal morphology, do not preserve verbal argument structure, and take adjectival modifiers.

3.4.2.2 Distribution of the nominalization in complement clauses

Nominalizations can occur with nearly all CTPs, except for modal verbs and the verb 'begin'. However, they often carry a nuance of meaning, such as manner (e.g. *kaft* dance:PTCP.PST 'dance, manner of dancing') or status as a cultural event (*k^{w}ævd* pray:PTCP.PST 'feast, prayer'), or they encode the semantic patient of the nominalized verb (*razərd* tell:PTCP.PST 'story, tale'). Such examples lie on the periphery of complementation.

3.4.3 Participial forms

The morphosyntax of the participial form in *-gɜ* (or *-gɜjɜ*, the ablative form of the participle) is described in detail by Belyaev and Vydrin (2011). It takes only one case marker, the ablative; possessive markers referring to the DO are possible, and the form in *-gɜ* may also (marginally) inflect for nominal number.

The participial forms take verbal morphological markers of aspect and negation. However, the finite negation (*nɜ*) can be replaced by the nominal preposition *ɜnɜ* 'without'. Tense and mood are not differentiated in these forms. The voice distinction is preserved; cf. (116) with the passive construction.

(116) *...mɜnɜ* *a-sə* *dəwwɜ* *tɜrx-ə* *=dɜr,* *iw* *ǯip:-ə* ***waʁd***
this.here this-ATTR two machine-GEN ADD one model-IN **issue.PTCP.PST**
wɜv-gɜ, *χisɜn* *kɜn-ənɜ* *kɜrɜzi-jɜ.*
be-CVBP separate do-PRS.3PL one.another-ABL
'Those two machines, which have been issued as the same model, differ from one another.'
(ONC)

The arguments of the participle are encoded in the same way as in the corresponding independent sentence.

In complementation, the participle form is used with verbs of immediate perception only:

(117) *ɜž* *fet:-on* *de=* *fšəmɜr-ə* *bɜχ-əl* *sɜw-**gɜ**.*
I see.PFV-PST.1SG thy brother-GEN horse-SPESPL go-**CVBP**
'I saw your brother riding a horse.'

When these verbs denote cognitive (indirect) perception, they cannot take the participle. In this case, the paratactic construction or the subordinators *kɜj* and *ɜmɜ* are used instead.

3.5 Conclusions

The Ossetic complementation system is comparatively rich: three non-finite strategies exist alongside a large number of subordinators available to introduce complement clauses. The finite strategies predominate. Non-finite strategies occur in the following contexts: the infinitive is restricted to control contexts with complements referring to the future (with respect to the temporal reference of the matrix clause); participles are used with immediate perception only; nominalizations are substantivized to a large extent and mostly denote not the situation itself, but institutionalized cultural events.

The number of complementation strategies is multiplied if we take into account the distribution of correlates. Correlative pronouns/adverbs are obligatory if the complement clause is preposed; otherwise they are used if the complement clause is presupposed, constitutes the topic, or encodes old/expected information. Hence, subordinators and correlative pronouns/adverbs bear different functions in Ossetic. The subordinator encodes the semantic type of the complement: *kɜj* (less commonly *ɜmɜ* or the paratactic construction) is used to introduce facts or propositions, *k^{w}əd* – events or caused situations, *k^{w}ə* – generic events (or stimulus of desire), *sɜmɜj* – propositions (mostly with reference to the future), *səma* – irrealis propositions, *kɜd* and *salənmɜ* – events with the verb 'wait'. The

correlate marks the status of the complement clause in the information structure of the sentence. An overview of the subordinators and correlative pronouns is presented in Table 1.

Table 1: Semantics of finite complementation strategies in Ossetic

Subordinator / type of construction	with a correlative pronoun	without a correlative pronoun
kəj	fact / topical proposition	proposition
əmə	topical / previously mentioned / expected proposition	proposition
paratactic construction	fact / topical proposition	proposition
kʷəd	previously mentioned or expected event or caused proposition	previously unmentioned or unexpected event or caused proposition
kʷə	previously mentioned or expected generic event (or stimulus of desire)	previously unmentioned or unexpected generic event (or stimulus of desire)
səməj	previously mentioned proposition, mostly with reference to the future	previously unmentioned proposition, mostly with reference to the future
səma	irrealis complement	–
kəd, *salənmə*	with the verb 'wait' only: event that is certain to happen	with the verb 'wait': event that is not certain to happen

The following conclusions can be drawn. The Ossetic system is sensitive to the opposition of coreferentiality patterns with verbs of causation, speech causation and potential action, thus demonstrating the same control pattern that is observed in SAE languages.

The most relevant semantic distinction in the Ossetic complementation system is that between events and facts/propositions. The 'presupposition vs. assertion' distinction is encoded by correlative pronouns/adverbs. However, the correlates are used to encode other semantic parameters, such as topicality/mentionedness/expectedness, and are not a subject to a strict grammatical rule (which brings this system close to the Russian complementation system, where the 'fact vs. proposition' distinction is mostly reflected in the intonation pattern or deduced from the context).

There is a special device for marking irrealis complements and generic events. An unexpected polysemy pattern is observed with the subordinator *kʷəd*, which can encode events and potential/caused situations in the future. Special devices are used with the verb 'wait'.

Another interesting feature is the use of relativization (the correlative construction) to encode facts and topical propositions. This brings the system of Ossetic close to the complementation systems found in North-West Caucasian languages and can be a result of areal influence (consult Belyaev & Serdobolskaya, in project, for details).

Appendix. Distribution of complementation strategies in Ossetic

Notation in the appendix (starting on the next page): «+» means that a construction is acceptable, «–» that it is unacceptable; «+/–» that variation exists among native speakers; «?» marks insufficient information. Note that adverbial clauses and indirect questions with *k^{w}əd*, *kɜd* etc. are not taken into account in the Appendix.

Complement-taking predicate	Infinitive in *-ən*	Nominalization in *-t-*	Participles in *-gɜ (-jɜ)*	*kɜj* 'that'	*kʷə* 'if, when'	*kʷəd* 'how', 'in order that'	*sɜmɜj* 'in order that'	*ɜmɜ* 'and'	*Parataxis*	*Other conjunctions*
Phasal verbs										
idɑjən 'begin'	+	–	–	–	–	–	–	–	–	
fɜwən 'finish, end'	–	+	–	–	–	–	–	–	–	
Modal predicates										
bon u 'can'	+	–	–	–	–	–	–	+	–	
žonən 'know, be able'	+	–	–	+	–	+	–	+	+	
fɜndən 'want'	+	–	–	–	+	+*[12]	+	+	+	
qɜwən 'must'	+	–	–	–	–	+*	+	?	+	
Predicates of emotion										
žɜrdɜmɜ sɜwən 'like'	+	+	–	+	+	+	–	+	–	
waržən 'love'	+	+	–	–	+	+	+	+	–	
tɜršən 'fear'	+	+	–	+	+	–	–	+	+/–	*žɜʁgɜ*
žɜrdɜ darən 'hope'	–	–	–	+	–	–	–	+	+	
χɜlɜg kɜnən 'envy'	–	+	–	+	–	+	–	–	–	
ɜχšəžgon u 'rejoice'	+	+	–	+	–	–	–	–	–	
sin kɜnən 'rejoice'	+	+	–	+	–	–	–	–	–	
diš kɜnən 'be surprised'	–	+	–	+	–	–	–	–	+	
Verbs of perception										
wənən 'see'	–	+	+	+	–	+	–	–	+	
qušən 'hear'	+/–	–	+	+	–	+	–	+	+	*žɜʁgɜ*
Mental predicates										
ɜnqɜlən 'think'	–	–	–	+	–	–	–	+	+	*səma*
ɜnqɜl wɜvən 'think'	–	–	–	+	–	–	–	+	+	*səma*
ɑftɜ kɜšən 'seem'	–	–	–	+	–	–	–	+	+	*səma*
wərnən 'believe'	–	–	–	+	–	–	–	+	–	

12 The asterisk means that the verb takes *kʷəd* in the meaning 'in.order.that'.

Complement-taking predicate	Infinitive in *-ən*	Nominalization in *-t-*	Participles in *-gɜ (-jɜ)*	*kɜj* 'that'	*kʷə* 'if, when'	*kʷəd* 'how', 'in order that'	*sɜmɜj* 'in order that'	*ɜmɜ* 'and'	*Parataxis*	*Other conjunctions*
ɜnqɜlmɜ kɜšən 'wait'	–	+	–	–	+	–	+	–	–	*kɜd, salənmɜ*
qʷədə kɜnən 'remember'	–	+	–	+	–	+	–	–	–	
roχ kɜnən 'forget'	+	+	–	+	–	+	–	+	+	
feroχ i 'forget'	+	+	–	+	–	+	–	–/+	+	
Speech verbs										
žɜʁən 'say'	+/–	+	–	+	–	+*	+	+	+	*žɜʁgɜ*
zurən 'tell'	–	+	–	+	–	+	+	-/+	–/+	
fɜršən 'ask'	–	+	–	–	–	–	–	–	+	wh-words, *žɜʁgɜ*
bar dɜttən 'let'	+	–	–	–	–	–	+	+	+	
wazən 'let'	+	–	–	–	–	–	–	+	–	
Predicates of potential action or causation										
žɜrdɜ ɜvɜrən 'promise'	+	–	–	+	–	–	–	+	+	
aχʷər kɜnən 'teach'	+	–	–	–	–	–	–	–	–	
aχʷər wən 'get used to'	+	+	–	+	–	–	–	–	–	
fɜlvarən 'try'	+	–	–	–	–	–	–	–	–	
arχajən 'seek'	+	+	–	–	–	–	+	+	–	
ɜχχwəš kɜnən 'help'	+	+	–	–	–	–	+	–/+	–	
qavən 'intend'	+	+	–	–	–	–	–	–	–/+	
χi sɜttɜ kɜnən 'prepare'	+/–	+	–	–	–	–	–	–	–	
ražə wɜvən 'agree'	+	–	–	+	–	–	+	–	–	
kɜnən 'make'	+	–	–	–	–	–	–	–	–	
Evaluative predicates										
χorž 'good'	+	+	–	+	+	–	–	+	–	
ɜvžɜr 'bad'	+	+	–	+	+	–	–	–	–	
žən 'hard'	+	+	–	–	+	–	–	–	–	

References

Abaev, Vassiliy I. 1950. *Grammaticheskij ocherk osetinskogo jazyka* [Grammar sketch of Ossetic]. Ordzhonikidze: Severo-Osetinskoe knizhnoe izdatelstvo.

Aissen, Judith L. 1974. Verb raising. *Linguistic Inquiry* 5(3). 325–366.

Apresyan, Juriy D. 1995. Problema faktivnosti: *znat'* i jego sinonimy [Problems of factivity: *znat'* and its synonyms]. *Voprosy jazykoznanija* 4. 43–64.

Arutjunova, Nina D. 1988. *Tipy jazykovyx znachenij: Ocenka. Sobytije. Fakt* [Types of linguistic meaning: Evaluation. Event. Fact]. Moscow: Nauka.

Asher, Nicholas. 1993. *Reference to abstract objects in discourse*. Dordrecht, Boston & London: Kluwer Academic Publishers.

Axvlediani, Georgiy S. (ed.). 1963. *Grammatika osetinskogo jazyka* [Ossetic grammar], vol. 1: *Fonetika i morfologija* [Phonetics and morphology]. Ordzhonikidze: NII Severo-Osetiskoj ASSR.

Bagaev, Nikolay K. 1965. *Sovremennyj osetinskij jazyk* [Modern Ossetic]. Part 1: *Fonetika i morfologija* [Phonetics and morphology]. Ordzhonikidze: Severo-Osetinskoe knizhnoe izdatelstvo.

Bagaev, Nikolay K. 1982. *Sovremennyj osetinskij jazyk* [Modern Ossetic]. Part 2: *Sintaksis* [Syntax]. Ordzonikidze: Severo-Osetinskoe knizhnoe izdatelstvo.

Belyaev, Oleg & Arseniy P. Vydrin. 2011. Participle-converbs in Iron Ossetic: Syntactic and semantic properties. In Agnes Korn, Geoffrey Haig, Simin Karimi & Pollet Samvelian (eds.), *Topics in Iranian Linguistics*, 117–134. Wiesbaden: Dr. Ludwig Reichert Verlag.

Belyaev, Oleg I. & Natalia V. Serdobolskaya. In project. Subordination in Ossetic and the typology of language contact: Syntactic replication without semantic influence.

Belyaev, Oleg I. 2011. Koreferentnost v korrelativax: svjazyvanie peremennoj ili anaphora? [Coreference in correlatives: variable binding or anaphora?]. *Acta Linguistica Petropolitana* 7(3). 261–270. Saint Petersburg: Nauka.

Belyaev, Oleg. 2014a. Systematic mismatches: Coordination and subordination at three levels of grammar. *Journal of Linguistics* 51. 267–326.

Belyaev, Oleg. 2014b. Anaphora in Ossetic correlatives and the typology of clause combining. In Pirkko Suihkonen & Lindsay Whaley (eds.), *On diversity and complexity of languages spoken in Europe and Northern and Central Asia* (Studies in Language Companion Series 164), 275–310. Amsterdam: John Benjamins.

Benveniste, Émile. 1966. *Problèmes de linguistique générale*, vol. 1. Paris: Gallimard.

Boye, Kasper. 2010a. Semantic maps and the identification of cross-linguistic generic categories: Evidentiality and its relation to Epistemic Modality. *Linguistic Discovery* 8(1). 4–22.

Boye, Kasper. 2010b. Reference and clausal perception-verb complements. *Linguistics* 48(2). 391–430.

Boye, Kasper. 2012. *Epistemic meaning: A crosslinguistic and functional-cognitive study* (Empirical Approaches to Language Typology 43). Berlin: Mouton de Gruyter.

COCA = The Corpus of Contemporary American English. http://corpus.byu.edu/coca/ (accessed June 2015).

Dalmas, Martine. 2013. Les corrélats de l'allemand et leur rôle dans la structure informationelle de l'énoncé. In Olga Inkova & Pascale Hadermann (eds.), *La corrélation. Aspects syntaxiques et sémantiques*, 307–321. Genève: Droz.

Dik, Simon C. 1997. *The theory of Functional Grammar*. Part 2: *Complex and derived constructions* (Functional Grammar Series 21). Ed. by Kees Hengeveld. Berlin & New York: Mouton de Gruyter.

Dixon Robert M. W. & Alexandra Y. Aikhenvald (eds.). 2006. *Complementation*. Oxford: Oxford University Press.

Gagkaev, Konstantin J. 1956. *Sintaksis osetinskogo jazyka* [Syntax of Ossetic]. Ordžonikidze: Severo-Osetinskoe knizhnoe izdatelstvo.

Givón, Talmy. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4(4). 333–377.

Green, Georgia M. 1976. Main clause phenomena in subordinate clauses. *Language* 52(2). 382–397.

Josephs, Lewis S. 1976. Complementation. In Masayoshi Shibatani (ed.), *Syntax and semantics* 5: *Japanese Generative Grammar*, 307–369. New York: Academic Press.

Karttunen, Lauri. 1971. Implicative verbs. *Language* 47(2). 340–358.

Kiparsky, Paul & Carol Kiparsky. 1971. Fact. In Leon Jakobovits & Danny Steinberg (eds.), *Semantics: An interdisciplinary reader*, 345–369. Cambridge: Cambridge University Press.

Krejdlin, Gregoriy E. 1983. O nekotoryx osobennostjax sintaksicheskogo povedenija predikatov s sentencialnymi aktantami [On some peculiarities of the syntactic properties of sentential complements]. *Semiotika i informatika* 21 [Semiotics and informatics], 76–88. Moscow: VINITI.

Kulaev, Nikolay X. 1959. *Sojuzy v sovremennom osetinskom jazyke* [Conjunctions in Modern Ossetic]. Ordzhonikidze: Severo-Osetinskoe knizhnoe izdatelstvo.

Miller, Philip. 2003. Negative complements in direct perception reports. In *Proceedings of the 39th Annual Meeting of the Chicago Linguistic Society* 1. 287–303.

Mithun, Marianne. 1999. *The languages of Native North America*. Cambridge: Cambridge University Press.

Nedjalkov, Vladimir P. 1979. Zametki po tipologii dvuxpredikatnyx konstrukcij. Opyt ischislenija [Notes on the typology of constructions with two predicates. Classification.]. In *Znachenije i smysl rechevyx obrazovanij* [The meaning and sense of pieces of discourse], 35–47. Kalinin: Kalinin State University.

Noonan, Michael. 1985. Complementation. In Shopen, Timothy (ed.), *Language typology and syntactic description* 2: *Complex constructions*, 42–140. Cambridge: Cambridge University Press.

Palmer, Frank R. 2001. *Mood and modality* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Peterson, Philip L. 1997. *Fact, proposition, event*. Dordrecht: Kluwer.

Podlesskaya, Vera I. 1990. Fakty, sobytija i propozicii v svete faktov japonskogo jazyka [Facts, events, and propositions in Japanese]. *Voprosy jazykoznanija* 4. 93–106.

Ransom, Evelyn N. 1986. *Complementation: Its meanings and forms* (Typological Studies in Language 10). Amsterdam & Philadelphia: John Benjamins.

Ross, John R. 1967. *Constraints on variables in syntax*. Cambridge: MIT dissertation.

Serdobolskaya, Natalia V. & Alexander V. Motlokhov. 2009. Semantika konstrukcij s sentencialnymi aktantami v adygejskom jazyke [Semantics of complementation in Adyghe]. In Yakov G. Testelec Pjotr M. Arkadjev, Alexander B. Letuchiy & Nina R. Sumbatova (eds.), *Aspekty polisintetizma: ocherki po grammatike adygejskogo jazyka*, 498–559. Moscow: RSUH.

Serdobolskaya, Natalia V. 2009. Semantika konstrukcij s sentencialnymi aktantami v kunkinskom variante darginskogo jazyka [Semantics of complementation in Qunqi Dargwa]. http://www.kibrik.ru/content/pdf/Serdobolskaja.pdf (accessed 25 March, 2009)

Serdobolskaya, Natalia V. 2011. K tipologii vyrazhenija znachenija genericheskogo sobytija v konstrukcijax s sentencialnymi aktantami [Towards a typology of means of expressing generic events in complementation]. *Acta Linguistica Petropolitana* 7 (3). 431–438. Saint-Petersburg: Nauka.

Serdobolskaya, Natalia V. This volume. Semantics of complementation in Adyghe.

Serdoboslkaya et al. = Serdobolskaya, Natalia V., Anfisa A. Iljevksya, Sergey A. Minor, Polina S. Miteva, Alexandra V. Faynveyc & Natalia S. Matveeva. 2012. Konstrukcii s sentencialnymi aktantami v finno-ugorskix jazykax [Complementation in Finno-Ugric languages]. In Ariadna I. Kuznetsova, Natalia V. Serdobolskaya, Svetlana Toldova, Sergey Saj & Elena Kalinina (eds.), *Finno-Ugorskie jazyki: fragmenty grammaticheskogo opisanija. Formalnyj i funkcionalnyj podxody* [Finno-Ugric languages: grammatical description. Formal and functional approaches], 382–476. Moscow: Jazyki slavjanskix kultur.

Suzuki, Satoko. 2000. Japanese complementizers: interactions between basic characteristics and contextual factors. *Journal of Pragmatics* 32. 1585–1621.

Verspoor, Marjolijn 2000. Iconicity in English complement constructions. In Kaoru Horie (ed.), *Complementation: cognitive and functional perspectives*. Amsterdam: John Benjamins.

Vydrin, Arseniy P. 2009. O leksichesko-sintaksicheskix sposobax vyrazhenija realnyx i irrealnyx znachenij (na materiale osetinskogo jazyka) [On lexico-syntactic means of expressing realis and irrealis (based on Ossetic data)]. *Izvestija Severo-Osetinskogo instituta gumanitarnyx i socialnyx issledovanij imeni V. I. Abaeva* 3(42). 171–182.

Vydrin, Arseniy P. 2011. *Sistema modalnosti osetinskogo jazyka v tipologicheskom osveschenii* [Typological description of the modality system of Ossetic]. Saint-Petersburg Institute for linguistic studies dissertation.

Vydrin, Arseniy P. forthcoming. Kosvennaya rech v osetinskom jazyke – forma *zhaergae* [Indirect speech in Ossetic – the form *zhaergae*]. *Acta Linguistica Petropolitana*. Saint-Petersburg: Nauka.

Xrakovskij, Viktor S. (ed.). 1985. *Tipologija konstrukcij s predikatnymi aktantami* [Typology of complementation]. Leningrad: Nauka.

Zaliznjak, Anna A. 1990. O ponjatii “fakt” v lingvisticheskoj semantike [On the concept of “fact” in semantics]. In Nina D. Arutjunova (ed.), *Logicheskij analiz jazyka. Protivorechivost i anomalnost teksta* [Logical analysis of language. Contradictions and anomalies of discourse.], 21–33. Moscow: Nauka.

Yaron Matras and Anton Tenser

Complementizers in Romani

1 Introduction

1.1 The language

Romani is the only Indo-Aryan language that has been spoken exclusively in Europe (and by European emigrants to the New World) since the Middle Ages. Historical reconstruction and analysis of loanword layers, coupled with the social-ethnographic profile of the speaker population and comparisons with similar Indian diaspora populations, suggest that the language may have been brought into Anatolia under Byzantine rule sometime around the eleventh century by a caste-like population specialising in itinerant services such as metalwork and entertainment. Today, Romani constitutes one of the largest minority languages of Europe, with upwards of 3.5 million speakers residing mainly in southeastern Europe (Romania, Bulgaria, Serbia, Macedonia, Greece, Albania) as well as in central Europe (Slovakia, Hungary). There are also sizeable Romani-speaking communities in the United States and in South America as well as in urban centres of western Europe. In the western fringe regions of Europe – Scandinavia, Britain, and the Iberian Peninsula – Romani populations abandoned their language toward the beginning of the nineteenth century or even earlier, but they maintain an in-group vocabulary of Romani origin, which typically comprises around 400–600 lexical roots. All Romani speakers are bilingual and often multi-lingual from a very young age. The language has traditionally been limited to oral usage, primarily within the extended family and with neighbouring clans. Codeswitching and language mixing are common and dialects of Romani have absorbed considerable structural and lexical influences from their respective contact languages.

The lexicon of most Romani dialects contains a shared inventory of only around 1,000–1,200 lexical roots. Of those, only around 800 are pre-European, consisting of a core vocabulary of around 600 Indo-Aryan roots along with layers of early loans from Iranian languages and from Armenian (as well as other Caucasian languages). All Romani dialects also contain a significant layer of Greek loans, as well as a much smaller number of shared lexical items of Balkan Slavic and Balkan Romance origin. The remainder of the vocabulary is typically derived from subsequent contact languages, and thus differs among the individual dialects of the language.

Romani morphology is partly fusional and partly agglutinative. The fusional element tends to be older and is best represented by the person conjugation on the verb, especially in the present tense, and by the retention of declension classes in the first layer of nominal inflection. The agglutinative element is generally younger and is best represented by second-layer case-markers, which are enclitic and not sensitive to declension class. Some tense and modality markers as well as verb derivational markers indicating transitivity and intransitivity are also agglutinative. On the whole the language can be said to be drifting toward stronger reliance on analytical structures. This is in part a contact phenomenon. The Romani dialects of western Europe have largely abandoned flexional derivations in the verb and rely instead on analytical constructions with auxiliaries. In southeastern Europe (Albania, Macedonia, and Greece), Romani dialects have developed analytical perfect tenses, while in central and eastern Europe mediopassive constructions tend to rely on the active verb with a reflexive pronoun. Inflectional case markers also show strong competition from prepositions, though the extent varies among the dialects. An outstanding characteristic feature of Romani morphology is the division between pre-European vocabulary and European loanwords in both nominal and verbal morphology. In nouns, the difference is expressed primarily in the choice of declension class, which determines the expression of the nominative and partly also the oblique form. In verbs, European loans require a loan-verb adaptation marker, which functions essentially as a derivational marker that mediates between the verb stem and its tense-aspect and person inflection.

Romani generally displays the properties of a SVO language with a flexible, pragmatically determined word order. Relative clauses follow their head nouns and contain resumptive pronouns wherever there is reference to the head in a non-subject role. Pro-dropping occurs especially in the chaining of subjects, and in this respect Romani is a pro-drop language to a limited extent. Clauses are generally finite and subordinations are introduced by conjunctions, which are typically derived from interrogatives or local relations expressions, or borrowed.

Geographical dispersion and the influence of diverse contact languages have led to the emergence of considerable difference among the dialects of Romani. A dense cluster of isoglosses, referred to as the Great Divide (Matras 2005), separates Romani dialects on both sides of what was, during the relevant period, the border zone (and war zone) between the Austrian and Ottoman empires. Once speakers adjust to a monolingual mode of discourse, it is generally possible for Romani speakers east of the Great Divide (from Greece to the Baltics) to understand one another, while the (much smaller) population of speakers to the west of the Great Divide speak dialects that are more fragmented.

Traditionally an oral language of the household, Romani is increasingly being used in electronic communication among community activists, who practice a pluralistic approach to codification and the selection of dialect features. Some programmes to promote the language in education and media are underway, though there is no standard written form of the language, nor a prestige variety that would lend itself for the creation of such a standard.

1.2 Methodological remarks

Functionalist typology is based on the assumption that structural configurations that are language-specific serve to encode communicative needs that are universal and which are grounded in cognitive processes of conceptualization. The methodological challenge that arises in this perspective is how to capture structural categories in a way that would do justice both to the way they are arranged in the individual language under discussion and in others that bear similarities to it (and in that respect may be said to share a 'type'), and to the universal conceptual meaning that they represent. Complements are generally understood to be propositions that are conceptually linked to a predicate, and which, structurally, are expressed at the sentential level as arguments of a verb (cf. Givón 1990: 515). The functionalist approach to complementation seeks to explore how the conceptual relationship between the verb and its complement proposition is linked to the structural relationship between them. Cristofaro (2003: 95–154) for instance, for a cross-linguistic sample of languages, maps a series of structural features (such as the presence/absence of TAM distinctions, case-marking, and person-agreement) to the semantic characteristics of predicates (such as modals, phasals, manipulative, desideratives, perception, utterance, and more). The cross-linguistic comparison yields implicational hierarchies of the likelihood of certain structural devices to correlate with certain semantic relations. Boye (2012) relates this kind of continuum to the notion of knowledge or 'epistemology': certain meanings require justification through explicit knowledge, or epistemic justification. Evidential and modal procedures provide various degrees of support for the transmission of knowledge. The intensity or strength of such support relates to the (universal) need to provide support. The structural system that provides the cues for conceptual support is the epistemic morphosyntactic system of a language. 'Epistemic' in Boye's (2012) terminology is thus the use of explicit (morphosyntactic) devices to provide support for those conceptual relations or elements of knowledge that require such support (epistemic or modal relations).

In the approach that we take below we follow a similar principle: we accept that structural devices convey instructions on how to process knowledge and

links between information chunks, including the conceptual links between verbs and their complements. We accept that such links are not of equal value, and that some require more (epistemic) support than others. We further accept that there is reason to anticipate a correlation between the distribution of structural devices that mark complementation, and the continuum of types of semantic relations, and in particular between the employment of explicit epistemic devices and the need to provide strong support. Since, however, we are dealing with dialects of a single language and thus with a limited set of morphosyntactic devices involved in complementation, we allow ourselves to focus on the principal parameters of the semantic-conceptual continuum that are of direct relevance to the category distinctions found in the relevant data. We therefore adopt Givón's (1990) notion of a continuum that represents the degree of 'semantic integration' between the two propositions in a complex clause (in our case: the main predicate and its complement). Semantic integration refers to the degree of independence of the two propositions, in terms of the cognitive processing of their content. In the centre of the semantic integration continuum are two main parameters. The first is the likelihood that the information conveyed by the complement is based on secure or real-world knowledge. Predicates of utterance and perception, for instance, are on the 'factual' side of the continuum: their complements are more likely to represent factual reality that can be confirmed independently of the event to which they are linked as complements. For this reason, they represent a lower degree of semantic integration, or a lower degree of dependency on the main predicate. Complements of phasal or desiderative – or non-factual – verbs, on the other hand, owe their factual realization to that of their main predicates (i.e. the likelihood of the phasal or desiderative action to yield the intended results). They are thus highly dependent on their main predicates, and so they represent tight semantic integration between main verb and complement. The second parameter is the degree of agentive control. Like factuality, agent control is a factor in the likelihood that the proposition conveyed by the complement will be realized. With desiderative verbs, for instance, agent identity across the two propositions is a strong guarantee for the likelihood of the action that is conveyed by the complement to be realized since it is dependent on the same conditions as the main predicate; semantic integration is therefore tight. By contrast, agent differentiation (manipulation) requires a weaker degree of control, representing a stronger potential for event independence and therefore weaker semantic integration. The aspect of control is not limited to complements. Predicates that are conceptually linked in a final or purpose construction offer a mixture of the two parameters in which independent predicates are linked through a non-factual relationship. For that reason, we include in our discussion also the occasional reference to the structure of purpose clauses in Romani.

1.3 An overview of complementizers in Romani

Complementation in this work is described separately from subordination following the structure of the Romani Morpho-Syntactic (RMS) database (http://romani.humanities.manchester.ac.uk/rms) and the typology of Romani complements along the factuality continuum, as proposed by Matras (1999: 18–20, 2002: 179–185, 2004). The two contrasting poles in this continuum are the non-factual predication (as in 'I want to go home') and factual predication (as in 'I know that he went home'), which together reflect the variation in the independent truth-value of the complement clause. The former (non-factual) accompanies modality verbs and is usually characterized in Romani by the complement particle *te*. The latter is best characterized as complementation with cognition-utterance verbs (cf. Givón 1990: 517). It accompanies factual complement phrases and is commonly expressed in Romani by the complementizer *kaj*. In the middle of this continuum lie the manipulation and purpose clauses, which are more ambiguous with respect to their independent truth-value, and for which the clause-linking strategies are subject to greater variation across dialects.

While the complementizer *te* is fairly stable across most Romani dialects, the functional slot that is represented in some dialects by inherited *kaj* is often filled by a borrowing from a contact language. The semantic functionality of the factual complementizer remains intact, however, even when the complementizer is replaced by a borrowing, and it is therefore customary in the context of Romani linguistics to refer to the *KAJ*-type complementizer (see Matras 2002, 2004). The distinction between *te* and *kaj* type complementation in Romani can be illustrated by the following examples (all data are taken from the Romani Morpho-Syntactic (RMS) database, based on original fieldwork carried out by the Manchester Romani Project: http://romani.humanities.manchester.ac.uk/rms. Sample numbers refer to RMS sample codes, detailed descriptions of which are fully accessible online, accompanied by audio files and transcriptions).

(1) *me kam-om [jekvar amerika **te** dž-au]*
I want-1SG once America **COMP** go-1SG
'I want to visit America someday.'
(Čuxny, Estonia, EST-005)

(2) *šundž-om [**kaj** inne romovja tiš ak beš-en]*
heard-1SG **COMP** other Roma also here live-3PL
'I heard that other Roma live here as well.'
(Bergitka, Poland, PL-007)

In the intermediate position of the independent truth-value continuum lie manipulation and purpose clauses. Their independent truth-value is more ambivalent,

and is dependent on the factor of agent control. Since control is a gradient, it is not surprising that we see greater variation in the forms of complementizers both across and within individual dialects. With the highest degree of agent control we often find the simple complementizer *te*; cf. example (3).

(3) *ov čjind-a nev-e furjavipe* [***te*** *šaj dž-al ando foro*]
he bought-3SG new-PL clothes **COMP** can go-3SG to town
'He bought new clothes so that he could go into town.'
(Gurbet, Croatia, HR-001)

With lower degrees of agent control we often find duplex or complex forms, where *te* appears alongside a 'reinforcer' or 'reinforcers' (Matras 2002: 181). One of the common reinforcers is in fact the factual complementizer *kaj*. Examples can be found in purpose and manipulation clauses:

(4) *phutregj-om e fereastra* [***kaj*** ***te*** *aštik alear-ap tut*]
opened-1SG ART window **COMP** **COMP** can hear-1SG you.OBL
'I opened the window so that I could hear you.'
(Keremidarea, Romania, RO-025)

(5) *mje manglj-em la* [***kaj*** *oj* ***the*** *žja-l tar*]
I asked-1SG her.OBL **COMP** she **COMP** go-3SG away
'I asked her to go away.'
(Kishinjovcy, Ukraine, UKR-007)

The present discussion will focus on the types of structures mentioned above in the context of the integration continuum, and will include modal, factual, and manipulation complements as well as occasional reference to purpose constructions. Indirect questions will not be discussed here, but for the sake of completeness it should be mentioned that in Romani these are realized through the use of the regular stock of interrogatives, without the use of complementizers.

(6) *phen man-ge* [***so*** *kerd-al adjejs*]
tell me-DAT **what** did-2SG today
'Tell me what you have done today!'
(Czech Vlax, Czech, CZ-001)

(7) *jou puht-as* [***kōn*** *jēl-o fōros*]
he asked-3SG **who** went-PTCP.M.SG town
'He asked who went to town.'
(East Finnish Romani, FIN-002)

There are no examples in the RMS database of the minimal pairs of the type "I know that…" vs. "I know whether…". Embeddings with the potential conditional

'whether' usually align themselves with modality (non-factual) clauses and are introduced by the non-factual complementizer *te*:

(8) *buč-om* *o* *profesar-es* [***te*** *v-el-a* *pe* *bijap*]
asked-1SG ART teacher-OBL **COMP** come-3SG-FUT to wedding
'I asked the teacher whether he will come to the wedding.'
(Sinti, Romania, RO-022)

The interpretation of the example in (8) as 'I asked the teacher *to* come to the wedding' is precluded by the use of the non-subjunctive form of the verb *v-* 'come'.

The position of an interrogative element in the complement is often occupied by loan elements from the contact languages. For example, dialects in contact with Russian often use the Russian interrogative-conditional particle *li*, those in contact with Polish often use the Polish *či*, and many dialects of the Balkans use the South Slavic *dali*:

(9) *me* *phučlj-om* *sykljaribnas-tar* [*av-ela* ***li*** *vov* *po* *bjav*]
I asked-1SG teacher-ABL come-3SG **COND** he to wedding
'I asked the teacher whether he will come to the wedding.'
(Servi, Ukraine, UKR-004)

(10) *phučl-om* *pes* *łe* *ranjen-dar* [***či*** *av-eła* *pe* *vera*]
asked-1SG REFL ART teachers-ABL **COND** come-3SG to wedding
'I asked the teacher whether he will come to the wedding.'
(Bergitka, Poland, PL-007)

(11) *phuč-um* *e* *nastavniko* [***dali*** *kam* *av-ol* *ko* *bijav*]
asked-1SG ART teacher **COND** FUT come-3SG to wedding
'I asked the teacher whether he will come to the wedding.'
(Arli, Serbia, YU-011)

An example of a minimal pair, where *te* and *kaj* might be used with a knowledge verb would be the non-factual *me džinav te rakirav* ('I know how to speak') vs. the factual *me džinav kaj jov rakirel* ('I know that he speaks').

Considering the dialect diversity of Romani and its geographical dispersion, we will mostly discuss tendencies rather than hard-set rules. The general trends will be exemplified using data samples from various dialects, in order to provide an appreciation of the diversity of Romani. The exceptions to the tendencies will be noted and exemplified. When a certain phenomenon is stated as a tendency and exemplified using data samples from only a few dialects, the reader should assume that the same phenomenon exists in most Romani dialects and that more extensive cross-dialectal exemplification is avoided due to the considerations of space.

1.4 Other aspects of complementation

1.4.1 Non-agreement

Romani is a predominantly finite language, in the sense that subordinate verbs agree with their subjects (Matras, 1999: 17). The overwhelming tendency across dialects is for the complement clauses to be finite, which is also a prominent feature of the Balkan languages. In the modal complement, the verb is marked for person and number:

(12) *me kam-av* [***te*** *sov-av*]
I want-1SG **COMP** sleep-1SG
'I want that I sleep.'
(General)

There are some exceptions, however. The neutralization of person agreement in complements has been referred to as 'the new infinitive' (Boretzky 1996). This, essentially 'de-balkanization' process (Matras 2002: 161) has affected the Sinti (German) and Finnish Romani dialects, the dialects of Poland, Bohemia and Slovakia, as well as the Romungro dialect of Hungary. The most frequently used 'infinitive' forms in Romani are the historical present tense personal markers: 2SG in *-(e)s* and 3SG in *-(e)l*:

(13) *jou pyrjylä* [***ceer-es*** *putti*]
he begins **do-INF** work
'He begins to work.'
(East Finnish Romani, FIN-008)

(14) *kamj-om* [*te* ***dža-ł*** *khere*]
wanted-1SG COMP **go-INF** home
'I wanted to go home.'
(Polish Xaladytka, Poland, PL-014)

(15) *ljubin-av* [*rano te* ***vypj-eł*** *jikh šjtekła kali*]
like-1SG morning COMP **drink-INF** INDF cup coffee
'I like to have a cup of coffee in the morning.'
(Polska Roma, Poland, PL-018)

(16) *na kam-av* [*te* ***dža-l*** *andro foros*]
NEG want-1SG COMP **go-INF** to town
'I do not want to go to town.'
(East Slovak, Slovakia, SK-002)

Eastern Romungro dialect of Slovakia uses a historical 2/3PL *-(e)n* form as the 'new infinitive':

(17) *na* *kam-es* [*te* ***dža-n***]
NEG want-2SG COMP **go-INF**
'You do not want to go.'
(Eastern Romungro, Slovakia, SK-027)

A remarkable development is found in the samples of the Ukrainian Servi and Xandžari speakers, where the infinitive form does not match any personal verb forms. Instead it looks like, on the surface, a present subjunctive form of the verb without the final consonant:

(18) *vov* *ačh-el* [*te* ***tire-Ø*** *buti*]
he begin-3SG COMP **do** work
'He begins to work.'
(Servi, Ukraine, UKR-004)

(19) *tu-te* *vydž-ala* [*te* ***dža-Ø*** *man-sa*]*?*
you-LOC can-3SG COMP **go** me-INS
'Can you come with me?'
(Kubanski Servi, Ukraine, UKR-008)

The most likely development path for this verb form is through the 2SG verb marker *-(e)h*, e.g. *tir-eh* → *tir-eØ*. Synchronically, it is clear that grammaticalization has been completed, since the 2SG marker in the Servi dialect is fully pronounced, and even velarized: *tir-ex* 'you do'. Thus, in this case one can speak of the emergence of a true infinitive form.

1.4.2 Tense, aspect, modality

Modal (non-factual) complement verbs in Romani almost always appear in the subjunctive:

(20) *me* *but* *mangl-em* [*te* ***ža-v*** *ande* *Indija*]
I very wanted-1SG COMP **go-1SG.SBJV** to India
'I wanted very much to go to India.'
(Sofia Erli, Bulgaria, BG-024)

There are some exceptions, however. The Romacilikanes dialect of Greece shows agreement between the modal verb and the complement verb in tense, aspect and modality – here using the remote past which has indicative imperfect/habitual reading in the given context:

(21) *panda* *kam-am-as* [*te* ***dža-v-as*** *ti* *Indija*]
always want-1SG-REM COMP **go-1SG-REM** to India
'I have always wanted to go to India.'
(Romacilikanes, Greece, GR-002)

In other Romani dialects, the above phrase shows the complement verb form in the subjunctive:

(22) *me uvek mang-av-as [te **dža-v** ki India]*
I always want-1SG-REM COMP **go-1SG.SBJV** to India
'I have always wanted to go to India.'
(Kovacki, Macedonia, MK-004)

With inherited modal verbs of ability, which are impersonal and do not carry tense markers, it is the verb of the complement clause that carries tense marking:

(23) *ame našy [Ø **arakhlj-am** rupun-i angrusti]*
we cannot Ø **found-1PL** silver-F ring
'We couldn't find the silver ring'
(Crimean, Ukraine, UKR-001)

(24) *me naštik [te **putard-em** o vudar]*
I cannot COMP **opened-1SG** ART door
'I couldn't open the door.'
(Gurbet, Serbia, YU-002)

A rare strategy to mark the tense of the action, found in the Ukrainian Gimpeny dialect sample, is to use the 3rd person past tense copula to form the past tense of the impersonal modal of ability:

(25) *me našty **sis** [te rakh-as e jangrusti]*
we cannot **were.3SG/PL** COMP find-1PL.SBJV ART ring
'We couldn't find the ring.'
(Gimpeny, Ukraine, UKR-020)

2 A descriptive account of Romani complementizers

2.1 Semantic aspects

Romani follows a general areal trend found in the languages of Balkans, where complements, including same-subject modal complements, are generally finite, and a semantic opposition between non-factual and factual complementizers is indicated through the identity of the complementizer. In Romani, the inherited and thus most widespread forms of the respective complementizers are *te* for the modal/non-factual, and *kaj* for the factual. Since the semantic-functional distinction tends to be maintained in the dialects even where the forms have undergone

structural change through substitution or borrowing, we can speak of complementizers of the type *TE* and *KAJ*, with individual formal representations in the individual dialects.

The inherited Romani modal complementizer *te* is realized in some dialects as *ti* (Sinti Romani) or *ta* (Lombard Italian Romani). The form is extremely stable across the dialects, and is rarely replaced by a borrowing. The one exception found in the samples is from the Xoraxani dialect, which borrows Bulgarian *da*:

(26) *bejkim* [***da*** *d-av* *ka-le* *romes-te* *duj-trin* *džigares*]
can.1SG **COMP** give-1SG this-OBL Gypsy.man-LOC two-three cigarettes
'I can give this Gypsy man some cigarettes.'
(Xoraxani, Bulgaria, BG-015)

The complementizer *te* is used in Romani consistently with the usual stock of modals: volition ('want'), positive and negative ability ('can', 'cannot'), inception and termination ('begin', 'start', 'end', 'finish'), obligation ('have to', 'need', 'must') and attempt ('try'), as well as other predicates, such as 'expect', 'like', 'fear', 'know how', 'dare' and 'have the strength' among others:

(27) *na* *kam-eha* [***te*** *dža-s* *ko* *čjüteti*]
NEG want-2SG **COMP** go-2SG.SBJV to town
'You do not want to go to town.'
(Mečkari, Albania, AL-001)

(28) *na* *mogindž-am* [*dava* *nikaj* ***te*** *rakh-as*]
NEG could-1PL this nowhere **COMP** find-1PL.SBJV
'We couldn't find it anywhere.'
(Polish Xaladytka, Poland, PL-014)

(29) *severim* [***te*** *pj-av* *ek* *čaša* *kafja* *sabajlen*]
like.1SG **COMP** drink-1SG INDF cup coffee morning
'I like to have a cup of coffee in the morning.'
(Kalajdži, Bulgaria, BG-009)

(30) *na* *has* *la* *zor* [***te*** *dža-l* *pale* *khere*]
NEG was.3SG her.OBL strength **COMP** go-3SG back home
'She did not have the strength to walk back home.'
(East Slovak, Slovakia, SK-011)

The emphasis of *te*-complements is on non-factuality, which is captured in the unconfirmed truth-value of the embedded proposition. Thus, even a verb that may, on the face of things, convey perception, but whose object cannot be confirmed in its factuality, will trigger the use of non-factual/modal *te* as the complement initiator:

(31) *džaker-ava* *m-e* *čhav-es* [***te*** *av-el* *kate* *sahati*]
wait-1SG my-OBL son-OBL **COMP** come-3SG here soon
'I expect my son to come here soon.'
(Romacilikanes, Greece, GR-002)

(32) *kam-au* [***ti*** *v-el* *mr* *čao* *kati* *an* *haki* *minuta*]
want-1SG **COMP** come-3SG my son here at any minute
'I expect my son to come here any minute.'
(Sinti, Romania, RO-022)

(33) *me* *adžukar-av* *mungr-e* *čhav-e* [***te*** *ares-el* *kate* *sa* *jekh*
I wait-1SG my-OBL son-OBL **COMP** return-3SG here any one
momento]
moment
'I expect my son to come here any minute.'
(Gurbet, Serbia, YU-002)

Co-acting with the factuality constraint is the constraint on agent control, which constitutes the second relevant semantic dimension on the continuum that represents Romani complementation. Semantic control is taken for granted in plain modality, of the kind illustrated above in examples (27) through (30). By contrast, we find variation in the structure of complementation when different degrees of semantic control appear. This is relevant to manipulation clauses, where the degree of control is a reflection of the power relations between the manipulator (the agent of the modal verb) and the manipulee (the agent of the embedded complement verb); see (34) through (42) and (5). It is equally relevant to purpose clauses, where control is reflected in the degree to which the intentional action (encoded in the final clause) is achievable through the action attributed to the agent in the main clause, in the case of same-subject purpose constructions as in (43) and (45), or indeed, again, in the power relations between the manipulator and manipulee, in the case of different-subject purpose constructions, as in (44).

While there is a continuum along the axis of agent control, there is also the independent factor of manipulative intent of the agent. If the manipulative intent of the agent is the focus of the proposition, complementation is not necessarily bound by the confines of the control continuum, and there is a tendency to express it through *te* across the dialects "irrespective of whether or not the target action is actually realized by the manipulee" (Matras 2002: 182):

(34) *mang-ava* [*akava* ***ti*** *naš-el*]
want-1SG he **COMP** leave-3SG
'I want him to go away.'
(Sofades, Greece, GR-004)

(35) *phend-em* *la-ke* [***te*** *čin-el* *purum* *aj* *šax*]
told-1SG her-DAT **COMP** buy-3SG onion and cabbage
'I told her to buy onions and cabbage.'
(Mexican Vlax, MX-001)

(36) *ou* *rodizj-a* *man* [***te*** *d-au* *les* *love*]
he asked-3SG me.OBL **COMP** give-1SG him money
'He asked me to give him money.'
(Ursari, Romania, RO-004)

Many dialects also show a strong tendency to use *te* with all manipulation clauses, even the ones with the strongest degree of control. This is true for samples from Albania, Serbia and Kosovo, most samples from Bulgaria, Croatia and Macedonia, samples from Hungary (Lovari and Gurvari), about half of the samples from Romania, as well as Crimean Romani samples from Russia and Ukraine, Romacilikanes dialect of Greece, and some Servi samples from Ukraine:

(37) *o* *dad* *kerd-a* *le* [***te*** *bičhav-el* *o* *lil*]
ART father made-3SG him.OBL **COMP** send-3SG ART letter
'His father made him send the letter.'
(Mečkari, Albania, AL-001)

(38) *me* *rod-em* *la-tar* [***te*** *dža-l-tar* *so* *maj* *sigo*]
I demanded-1SG her-ABL **COMP** go-3SG-away what very quickly
'I demanded for her to leave immediately.'
(Gurbet, Serbia, YU-002)

(39) *oj* *nateringj-a* *le* [***te*** *dža-l* *pes*]
she made-3SG him.OBL **COMP** go-3SG REFL
'She made him leave.'
(Arli, Macedonia, MK-002)

(40) *bar-e* *manuša* *thejard-e* [***te*** *thar-as* *amar-e* *khera*]
big-PL men made-3PL **COMP** burn-1PL our-PL houses
'The government made us burn our houses.'
(Crimean, Ukraine, UKR-001)

(41) *oj* *kerd-as* *os* [***te*** *naš-el*]
she made-3SG him.CL **COMP** leave-3SG
'She made him leave.'
(Romacilikanes, Greece, GR-002)

(42) *voj* *muk-el* *e* *farfurja* [***te*** *pher-el*]
she let-3SG ART plate **COMP** fall-3SG
'She lets the plate fall.'
(Laeši Kurteja, Moldova, MD-007)

The control and integration continuum with purpose clauses is manifested in Romani primarily in terms of how (un)contentiously the action of the main clause will lead to the outcome of the complement clause. The more certain the outcome, the more integrated the event is, and thus the more likely we are to see the simplex complementizer *te*. The less certain the outcome, the less control the agent is perceived to have over it, the more likely we are to see the more complex, multi-element complementizers (see below). The weakest degree of control can be seen with potential purpose propositions, as in 'in order to A, one has to B'. Out of the available samples, the continuum is best illustrated with data from Goli Cigani dialect of Bulgaria, where we see, in order of decreasing integration, *te* vs. *za te* vs. *za da te*:

(43) *avij-om* *khere* [***te*** *dikh-av tut*]
came-1SG home **COMP** see-1SG you.OBL
'I came home to see you.'
(Goli Cigani, Bulgaria, BG-011)

(44) *oj* *veče* *dij-as* *t-e* *phales-te* *o* *paras* [***za*** ***te***
she already gave-3SG your-OBL brother-LOC ART money **COMP** **COMP**
dža-l *maškare*]
go-3SG to.the.middle
'She already gave your brother the money to go to town.'
(Goli Cigani, Bulgaria, BG-011)

(45) *oj* *tho-ela* *o* *xurd-es* *te* *skemba* [*te* *beš-el*] [***za*** ***da***
she put-3SG ART child-OBL on chair COMP sit-3SG **COMP** **COMP**
te *parvar-el* *les*]
COMP feed-3SG him.OBL
'She sits the child down on the chair to feed him.'
(Goli Cigani, Bulgaria, BG-011)

While Goli Cigani shows the full spectrum of the continuum, many dialects show a uniform use of simplex *te* with all purpose clauses. Many of these are the same samples that use exclusively simplex *te* with manipulation clauses, as discussed above. We find this non-differentiation along the control continuum with all samples from Croatia, Macedonia, Serbia and Kosovo, both Hungarian Lovari and Gurvari, many dialects of Romania and Moldova (plus Vlax dialects elsewhere: Russian Lovari, Czech Vlax and Ukrainian Shanxajcy), some samples from Bulgaria (Xoraxani), as well as Crimean Romani dialect. All of these dialects show simplex *te* even with the weakest degree of control:

(46) *oj* *pučlj-a* *ma* *so* *te* *čjer-el* [***te*** *av* *la*
she asked-3SG me.OBL what COMP do-3SG **COMP** come her.OBL
majbut *love*]
more money
'She asked me what to do to earn some more money.'
(Gurbet, Croatia, HR-001)

(47) *d-em* *e* *gajž-es* *love* [***te*** *kin-el* *kaveja*]
gave-1SG ART man-OBL money **COMP** buy-3SG coffee
'I gave the non-Gypsy man some money so that he could buy coffee.'
(Lovari, Hungary, HU-004)

(48) [***te*** *džja-n* *ke* *djukjana*] *tumen-ge* *kam-el* [*te* *džja-n* *karjin* *ke*
COMP go-3PL to shop you.PL-DAT want-3SG COMP go-3PL toward to
khangerji]
church
'To go to the shop, you have to walk towards the church.'
(Crimean, Russia, RUS-011)

The *KAJ*-type complementizer in Romani (as mentioned above, we use the archetype *KAJ* to capture the fact that this function is often filled by borrowed forms) occurs with various cognition-utterance verbs, to use the term from Givón (1990: 517), such as 'see', 'hear', 'say', 'know', 'understand', 'think', 'feel', 'seem' and 'be certain', among others. There are no co-reference restrictions between arguments of the main and the complement clause. The argument of the complement clause can refer either to the same subject as the main clause, or to a different subject:

(49) *voj* *phend-a* [***kaj*** *či* *pindžard-a* *khanikaj*]
she said-3SG **COMP** NEG knew-3SG nobody
'She said that she did not know anyone.'
(Gurbet, Croatia, HR-001)

(50) *me* *šundž-om* [***kaj*** *jov* *adaj* *džid-o* *ot* *červca*]
I heard-1SG **COMP** he here live-PTCP.M.SG from June
'I heard he has lived here since June.'
(Polish Xaladytka, Poland, PL-014)

(51) *me* *džan-au* [***kaj*** *sas* *tu* *but* *bučji* *gadalaj* *duj* *brš*]
I know-1SG **COMP** was you much work these two years
'I know that you had a lot of work during the past two years.'
(Manuša Čurjaja, Croatia, HR-003)

(52) *me* *dedum-av* [***kaj*** *kažno* *manuš* *gindisar-el* *kajci* *pal* *pe*]
I reckon-1SG **COMP** every man think-3SG only of REFL
'It seems to me everybody thinks only of themselves'
(Xandžari, Ukraine, UKR-010)

While these examples show how the prototypical function of *KAJ* is related to factuality (i.e. potentially confirmed truth-value of the embedded proposition), *KAJ* may also combine with *te* to form a duplex complementizer. Typically, this strategy of using *KAJ* as a 'reinforcer' for the non-factual complementizer *te* is found in cases of weak agent control, often in cases of different-subject manipulation and purpose clauses, or more generally, with weaker semantic integration of the two clauses, the main and embedded clause:

(53) *zaphand-av* *e* *blaka* [***kaj*** *amen* ***te*** *na* *šun-el*]
shut-1SG ART window **COMP** us.OBL **COMP** NEG hear-3SG
'I will shut the window so that (he) can't hear us.'
(East Slovak, SK-002)

(54) *jej* *čuv-el* *bajatos* *pale* *skamint* [***kaj*** ***te*** *d-el* *te*
she put-3SG child.OBL on table **COMP** **COMP** give-3SG COMP
xa-l *les*]
eat-3SG him.OBL
'She seats the kid on the chair to feed it.'
(Plasčuny, Ukraine, UKR-019)

(55) *avilj-om* *khere* [***kaj*** ***te*** *dikh-av* *tut*]
came-1SG home **COMP** **COMP** see-1SG you.OBL
'I came home to see you.'
(Romungro, Romania, RO-059)

(56) *ov* *kind-as* *neve* *cavala* [***ka*** ***te*** *dža-l* *ti* *poli*]
he bought-3SG new.PL clothes **COMP** **COMP** go-3SG to town
'He bought new clothes so that he could go into town.'
(Romacilikanes, Greece, GR-002)

(57) *vov* *phendj-a* [***kaj*** *me* ***te*** *bikin-au* *e* *tradini*]
he told-3SG **COMP** I **COMP** sell-1SG ART car
'He told me to sell the car.'
(Laeši Kurteja, Moldova, MD-007)

(58) *kam-av* [***kaj*** ***te*** *dža-l* *het*]
want-1SG **COMP** **COMP** go-3SG away
'I want him to go away.'
(East Slovak, SK-002)

The overall tendency across Romani dialects is to use simplex *te* with tighter integrated events, and to use a duplex or complex linking elements for those events that are less tightly integrated. In cases where a duplex complementizer is used, one of the elements is usually *te*, while the other, reinforcing element is usually the *KAJ*-type complementizer used by the particular dialect. With more complex linking, the additional reinforcing elements often come from borrowed particles

(conjunctions and prepositions) that are used in similar constructions in the contact language(s). Matras (2002: 182) suggests that this demonstrates two-level iconicity: tighter semantic integration correlates to a) simpler form of the subordinator (in the case of *te* complements), and b) to inherited forms (both *kaj* and *te*), while less tight integration is associated with more complex subordination strategies and is much more susceptible to influence from the outside language. Thus, Lombard Sinti of Italy uses duplex *par te*, where the first element is the borrowed Italian purpose particle. Romungro dialect of Hungary uses duplex *hodj te*:

(59) *pirad-om* *i* *vali* [***par*** ***ta*** *sun-a-to*]
opened-1SG ART window **COMP** **COMP** hear-1SG-you
'I opened the window so that I can hear you.'
(Lombard Sinti, Italy, IT-011)

(60) *khēre* *āj-om* [***hodj*** ***te*** *dikh-av tut*]
home came-1SG **COMP** **COMP** see-1SG you.OBL
'I came home to see you.'
(Romungro, Hungary, HU-009)

Many dialect samples use the opposition of *te* vs. borrowed *KAJ*-type reinforcer + *te* in both manipulation and purpose clauses. This is true of the dialects with strong Russian influence, which use *so(b)* as the reinforcer. Sofades dialect of Greece uses Greek loan *ja* as a reinforcer in both types of clauses. Some East Slovak samples use Hungarian-loan *hoj*, while many Finnish Romani samples often utilize either Swedish *at* or Finnish *et(tä)*. In cases where the strategies of manipulation clause and purpose clause linking differ, the purpose clause strategies seem to be more complex and specialized, and often involve more specific borrowing from the other languages, while manipulation clause linking strategies are more likely to match in form the *KAJ*-type particles found with factual complementation. Examples of more specialized purpose clause complements that are borrowed include Italian *par* in *par ta*, Slavic *bi* in *ta bi ta*, South Slavic *za* and *da* in *za da te*, while examples of more complex Romani-internal developments for purpose clause linking are Finnish Romani *toolesko khaal*, and *kaš (de) te* found in several Vlax-type samples (see Section 2.5 below on combinability of complementizer elements).

Table 1 summarises the effect of semantic integration on the choice of complementizer in Romani. Tighter semantic integration (higher agent control and greater likelihood of the success of the main predication to entail the realisation of the proposition contained in the embedded predication) is likely to correlate with the use of a simplex complementizer, which is almost invariably the inherited form *te*. By contrast, loose semantic integration (weaker control of the agent of the main predication over the agent of the embedded predication, or weaker

likelihood of realisation of the proposition encoded by the embedded predication) is more likely to correlate with the use of a complex form.

Table 1: Manipulation and purpose clauses – tendencies

INTEGRATION:	tighter		→		looser
COMPLEXITY:	simplex	→	duplex	→	complex
SOURCE:	inherited		→		borrowed
FORMS:	*te*	→	+ *KAJ*-type	→	+ other

2.2 Distribution

Romani complementizers appear with modal (non-factual) and factual predicates, as well as in purpose and manipulation clauses. The inherited modal complementizer *te* is stable and found in most of the dialects with all modals, including volition ('want'), positive and negative ability ('can', 'cannot'), inception and termination ('begin', 'start', 'end', 'finish'), obligation ('have to', 'need', 'must') and attempt ('try'), as well as other predicates, such as 'expect', 'like', 'fear', 'know how', 'dare' and 'have the strength'. It is also used regularly across dialects in manipulative and purpose clauses. Factual complementation is accomplished through the use of the inherited complementizer *kaj*, or through borrowed complementizers. The *KAJ*-type complementizer appears with various factual cognition-utterance verbs, such as 'see', 'hear', 'say', 'know', 'understand', 'think', 'feel', 'seem' and 'be certain', among others. In addition, it often appears as a reinforcer with manipulation clauses, including desiderative predicates ('want', 'wish'), manipulating predicates ('ask', 'demand', 'tell'), as well as with purpose clauses. In addition various borrowed and calqued markers are used with factual, and especially manipulation and purpose complements. Table 2 gives examples of the distribution of complementizer types across different main clause predicates from four distinct Romani dialects.

In the table below, Mečkari represents the simplest system, where only two forms *-te* and *kaj-* are in binary opposition: *kaj* is found with factual complementation, whereas *te* is used with modal, manipulation and purpose clause predicates. This binary opposition is found in most samples from Croatia, Macedonia, Serbia and Kosovo, many of the samples from Bulgaria and Romania, as well as Hungarian Lovari and Gurvari. This is also true for Crimean Romani, and Šanxajcy and Xandžari samples from Ukraine.

Table 2: Complementizers with various predicates

predicate	Mečkari, Albania (AL-001)	Bergitka, Poland (PL-007)	Kalderaš, Romania (RO-065)	Gimpeny, Ukraine (UKR-020)
epistemic				
see	n/a	*kaj*	*kaj*/*kə*	*kaj*
hear	*ka*(*j*)	*kaj*	*kaj*	*kaj*
say	*ka*(*j*)	*kaj*	*kaj*/*kə*	*kaj*
know	*ka*(*j*)	*kaj*	*kaj*	*kaj*
understand	*ka*(*j*)	*kaj*	*kaj*	*kaj*
think	*ka*(*j*)	*kaj*	*kaj*/*kə*	*kaj*
feel	n/a	*kaj*	*kaj*/*kə*	*kaj*
modality				
want	*te*	*te*	*te*	te
can	*te*	Ø	*te*	te
cannot	*te*	Ø	*te*	te
must	*te*	*te*	*te*	te
begin	*te*	*te*	*te*	te
try	*te*	n/a	n/a	te
like	*te*	*te*	*te*	te
fear	*te*	*te*	*te*	te
know how to	*te*	*te*	n/a	te
dare	*te*	*te*	*te*	te
manipulation				
want	*te*	*kaj te*	*ka*(*j*) *te*	*kaj te*
demand	*te*	*kaj te*	*te*	*kaj te*
ask	*te*	*kaj te*	*te*	*kaj te*
tell	*te*	*kaj te*	*te*	*te*
convince	*te*	*kaj te*	*te*	*te*
let	*te*	*kaj te*	*te*	*te*
independent clause				
purpose clause	*te*	*kaj te*	*kaš te*, *te*	*te*, *kaj te*

Many dialect samples show a three-way distinction, with the usual modal *te* and factual *kaj* (or *KAJ*-type), and a duplex complementizer *kaj te* used consistently with both manipulation and purpose clauses. This three way distinction with inherited *kaj*, exemplified with Bergitka in the table above, is found in many dialect samples from Slovakia, some samples from Romania (Romungro, Spoitori, Ursari), all of the samples from Poland (Polska Roma, Bergitka, Polish Xaladytka), and several samples from Ukraine (Plasčuny, Kišinjovcy). The three way distinction where a borrowed or calqued element is used instead of *kaj* is

found in many samples that are in strong contact with Russian, including dialects from Russia, Lithuania, Latvia, and the Servy dialects of Ukraine.

Romanian Kalderaš in the table above is one of the few dialects of Romani where manipulation and purpose clauses show different complementation strategies. In all such cases, the strategies found in purpose clauses are more complex and specialized than those found in manipulation clauses. Other such dialects include Lombard Sinti, which uses *par ta* in purpose clauses; Italian Molise, which uses the Slavic irrealis particle (*ta bi ta*); some dialects from Bulgaria, which use complementizer elements borrowed from South Slavic (*ta te*, *za da te*); and some of the samples from Finland, which use a language-internal derivation *toolesko*, a demonstrative pronoun 'there' with a genitive case ending.

Finally, the Gimpeny sample in Table 2 illustrates that some of the dialects do not have a clear-cut split in the purpose and manipulation clauses; rather, the simpler form *te* is used when there is a tighter integration between the predicate and the complement clause, and more complex strategies are used when there is looser integration.

2.3 Complementizers as a system

The Romani system of complementizers can be discussed at three levels: functional-semantic, structural, and diachronic (etymological). At the functional-semantic level, the principal poles are factuality and non-factuality, with predicates and predicate combinations arranged, potentially, on a continuum in between the two, and continuum points being sensitive to the likelihood of realization of the embedded predicate, which entails the degree of semantic control of the main agent over the action depicted by the embedded predicate, either directly (over that action itself) or indirectly (through control of a secondary agent or manipulee). At the structural level, the continuum shows a range from the absence of complementizers (see the discussion in Section 2.4 below), through the use of simplex complementizers, until the use of duplex and complex complementizers. By and large, the structural continuum maps directly onto the functional-semantic continuum (see Table 1). As a language that is permanently in contact, and with dispersed dialects showing a history of contact with different languages, Romani offers a unique laboratory for inquiries into contact-induced structural change. Etymology thus figures prominently in the area of complementizers, which are often prone to borrowing, though under certain constraints, which often equally map onto the semantic-functional continuum (see also discussion in Elšík & Matras 2009). Table 3 summarizes all the complementizers commonly found across the dialects, including the types of predicates that they

are associated with, their etymology, and the additional functions of the complementizer forms, where appropriate (for a discussion of additional semantic functions see Section 2.6 below).

Table 3: Complementizer forms, functions and etymology

complementizer (including simplex and duplex)	type of predicate (semantics)	etymology (inherited/borrowed)	additional functions
te	modals, purpose, manipulation	inherited	conditional, optative/imperative
Ø	inherited modals of (in) ability		
kaj	factual	inherited	interrogative 'where', relativizer
ke, *oti*, *hod*/*hoj*, *ani*	factual	borrowed	
so, *sy*(*r*)	factual	calqued	interrogative
kaj te	manipulation, purpose	inherited	
et(*tä*), *ta*	factual, manipulation, purpose		
ke te, *hod*/*hoj te*, *ja ti*	manipulation, purpose	borrowed + inherited	
sob(*y*), *sob*(*y*) *te*	manipulation, purpose	semi-calqued	conditional *by*
za te, *za da te*, *par te*, *ta te*, *kaš te*	purpose	borrowed + inherited	
toolesko (*khaal*)	purpose	language-internal	genitive case

2.4 Complementizer omission

The general tendency across Romani is to use the overt complementizer *te* for modal complementation. The form is very stable, both in terms of being highly immune to replacement through borrowing, and in terms of covering the usual stock of modals. There are several contexts, however, where this complementizer is "omitted" (we prefer to speak simply of the absence of a complementizer, rather than its "omission"). Absence of a complementizer (of the *KAJ*-type) with factual complements is rare and can be attributed to a paratactic-like structure, where the main predicate is usually separated from the embedded clause by a pause, or

a re-formulation or sequential (rather than embedded) structure can at least be inferred:

(61) *dikh-ava* [**Ø** *savre* *düšünüiler* *sāde* *pumen-ge*]
see-1SG **Ø** everybody thinks only REFL.PL-DAT
'Everybody seems to be thinking only of themselves.'
(Xoraxani, Bulgaria, BG-023)

(62) *šund-an* [**Ø** *vi* *aver* *rom* *beš-en* *kote*]
heard-2SG **Ø** also other Roma live-3PL here
'You've heard that other Roma live here as well.'
(Lovari, Hungary, HU-004)

Elsewhere, zero-complementizer is restricted to modality clauses:

(63) *me* *kam-am* [**Ø** *nodža-l* *pes-ke* *jou*]
I want-1SG **Ø** leave-3SG REFL-DAT he
'I want him to go away.'
(Čuxny, Latvia, LV-005)

Here too, however, it is often subject to variation; cf. (64) and (65).

(64) *ame* ***našy*** [**Ø** *arakhlj-am* *rupun-i* *angrusti*]
we **cannot** **Ø** found-1PL silver-F ring
'We couldn't find the silver ring.'
(Crimean, Ukraine, UKR-001)

(65) *ov* *ni* ***može-t*** [***te*** *dža-l*]
he NEG **can-3SG.BOR** **COMP** go-3SG
'He can't leave.'
(Crimean, Ukraine, UKR-001)

The generalisation of zero-complementizer as a rule is bound to certain modality predicates, which are arranged on a hierarchical continuum (cf. Elšík & Matras 2009). The Ø complementizer is commonly found with the inherited modals of positive and negative ability ('can' and 'cannot'), both of which are impersonal (not inflected for person and number). Out of the two, the modal of negative ability (*našti(k) / naši)* is diachronically much more stable across the dialects, while the modal of positive ability is much more prone to borrowing from the contact languages. The inherited form of this latter modal, *ašti(k)/sašti/hašti/vašti/šaj*, is replaced in various Romani dialects with Slavonic *mog-*/*mož-*, Greek *bor-*, Italian *pot-*. As a rule, the inherited forms of the modals of ability are more likely to take a zero-complementizer, cf. examples (66)–(69), while the borrowed forms take the usual modal complementizer *te*, cf. examples (70)–(71).

(66) ***naši*** [**Ø** *sikjav-a*] *soskətu* *trjabva* [*tə* *pomeizə-a* *m-ə* *da-ke*]
cannot **Ø** study-1SG because needed COMP help-1SG my-OBL mother-DAT
'I cannot study because I have to help my mother.'
(Muzikantska Roma, Bulgaria, BG-010)

(67) ***nasti*** [**Ø** *putr-av-as* *i* *porta*]
cannot **Ø** open-1SG-REM ART door
'I couldn't open the door.'
(Sofades, Greece, GR-004)

(68) *me* ***ašti*** [**Ø** *d-av* *les* *xari* *cigares*]
I **can** **Ø** give-1SG him.OBL some cigarettes
'I can give him some cigarettes.'
(Sofia Erli, Bulgaria, BG-024)

(69) *ov* *či* *darajlo* *kaj* ***ša*** [**Ø** *per-o*]
he NEG scared.M COMP **can** **Ø** fall-PTCP.M.SG
'He wasn't afraid that he might fall.'
(Gurbet, Hungary, HR-001)

(70) *tu* ***možyn-es*** [***te*** *zasuv-es* *txaves-a* *da* *ljoxi*]*?*
you **can-2SG** **COMP** sew-2SG thread-INS these holes
'Can you mend these holes with thread?'
(Lithuanian Romani, LT-007)

(71) ***borin-eha*** [***ti*** *pus-es* *les*]
can-2SG **COMP** believe-2SG.SBJV him.OBL
'You can believe him.'
(Sofades, Greece, GR-004)

There are exceptions to this rule, however. Many dialects of the Southern Balkans, while keeping the inherited forms of the modals of ability show a tendency to use the overt complementizer *te*:

(72) *ame* ***našti*** [***te*** *irin-as* *amen*]
we **cannot** **COMP** return-1PL REFL.1PL
'We cannot go back.'
(Sofia Erli, Bulgaria, BG-024)

(73) *mislisar-av* *kaj* ***šaj*** [*theara* **te** *iriv* *kir-e* *pare*]
think-1SG COMP **can** tomorrow **COMP** return your-PL money
'I think that I will be able to pay you back tomorrow.'
(Gurbet, Macedonia, MK-001)

(74) ***šaj*** [***te*** *pomožin-av* *tumen-gje*]
can **COMP** help-1SG you.PL-DAT
'I can help you to.'
(Arli, Macedonia, MK-002)

(75) ***našti*** [***te*** *irin* *amen*]
cannot **COMP** return REFL.1PL
'We cannot go back.'
(Arli, Macedonia, MK-002)

Exceptions of the opposite kind – borrowed modals of ability with zero-complementizer – are also found, usually with borrowed modals that are impersonal:

(76) ***može*** [**Ø** *inanasən-əs* *les-ke*]
possible **Ø** believe-2SG him-DAT
'(You) can believe him.'
(Xoraxani, Bulgaria, BG-015)

(77) *nam* ***pot*** [**Ø** *studin-a* *pe*]
NEG **can** **Ø** study-1SG REFL
'I cannot study.'
(Molise, Italy, IT-010)

(78) *ni* ***miga*** [**Ø** *v-as* *pali*] *finke* *joj* *na* *sasto-la*
NEG **can** **Ø** come-1PL back until she NEG become.healthy-3SG
'We cannot go back until she gets well.'
(Lombard, Italy, IT-011)

When the borrowed modal has personal forms, however, the tendency is to use an overt complementizer *te*:

(79) ***bori-s*** [***te*** *ker-es* *ola* *gures* *šeles-a*]*?*
can-2SG **COMP** do-2SG these holes thread-INS
'Can you mend these holes with thread?'
(Romacilikanes, Greece, GR-002)

(80) *jesli* *tumə* *kam-ən* *mə* ***možyn-ou*** [*tumən-ge* ***te*** *pomožyn-ou*]
if you.PL want-2PL I **can-1SG** you.PL-DAT **COMP** help-1SG
'If you (pl) want I can help you.'
(Lithuanian Romani, LT-005)

There are also cases when the inherited modals of ability become personal in individual dialects, and inflect for person and number. In these cases, the overt complementizer *te* is always used:

(81) ***naštisar-as*** [***te*** *bold-as*]
cannot-1PL **COMP** return.1PL
'We cannot go back.'
(Mexican Vlax, Mexico, MX-001)

(82) ***dašti-s*** [***te*** *av-es* *man-ca*]*?*
can-2SG **COMP** come-2SG me-INS
'Can you come with me?'
(Mexican Vlax, Mexico, MX-001)

(83) ***nasčind-em*** [***te*** *pətr-av* *o* *udar*]
could.not-1SG **COMP** open-1SG ART door
'I couldn't open the door.'
(Piculesti, Romania, RO-013)

(84) *či* ***šajnd-em*** [***te*** *arak-au* *les* *ničі-sar*]
NEG **could-1SG** **COMP** find-1SG him.OBL NEG-how
'I couldn't find it anywhere.'
(Kurturare, Romania, RO-015)

(85) ***dašti-l*** [***te*** *d-el* *bryšynd* *i* *ando* *Julio*]
can-3SG **COMP** give-3SG rain is in July
'It is possible that it will rain in July.'
(Rakarengo, Romania, RO-002)

The same phenomenon exists in various dialects of Moldova, as well as in those dialects of Ukraine that have arrived there relatively recently, in the last 150 years. This seems to be an areal phenomenon, affecting different dialects with current and recent contact with the Romanian language.

Sometimes the inherited modals of ability are replaced through Romani-internal material, as in the case of the Čuxny (Estonian Romani) verb *dole-* (< Russian aktionsart *do-* + Romani verb *le-* 'take'). Another example is the Kubanski Servi *vydža-* (< Russian aktionsart *vy-* + Romani verb *dža-* 'go'), which is a calque from the Russian *vy-xodit'* 'go out, come out'. In these cases the modal is inflected for person, and, as expected, it takes an overt complementizer:

(86) ***dol-esa*** *tu* [*man-ca* ***te*** *j-ēs*]*?*
can-2SG you me-INS **COMP** come-2SG
'Can you come with me?'
(Čuxny, Estonia, EST-005)

(87) *tu-te* ***vydža-la*** [***te*** *dža* *man-sa*]*?*
you-LOC **can-3SG** **COMP** go me-INS
'Can you come with me?'
(Kubanski Servi, Ukraine, UKR-008)

While it is quite common to find zero-complementizers with the impersonal modals of positive and negative ability, all other modals in Romani are usually personal and necessarily take the overt complementizer *te*. Exceptions are Finnish

Romani and Lotfitka Romani of Latvia; both of these dialects show optional zero realization of modal complementizers:

(88) *tu* *mostul-as* [**Ø** *aav-en* *man-go*]
you must-REM **Ø** come-2PL me.DAT
'You should visit me.'
(East Finnish Romani, FIN-008)

(89) *mir-i* *tykn-i* *čajori* *straxadžo-la* [**Ø** *dža-l* *pirdal* *phurt*]
my-F little-F daughter fear-3SG **Ø** go-3SG across bridge
'My little daughter is scared to go across a bridge.'
(Lotfitka, Latvia, LV-005)

(90) *tu* *na* *kam-es* [**Ø** *dža-s* *po* *foros*]
you NEG want-2SG **Ø** go-2SG to town
'You do not want to go to town.'
(Lotfitka, Latvia, LV-006)

As with the modals of ability, zero realization of the complementizer is more likely to appear in impersonal modal constructions. East Slovak Romani has the zero modal complementizer with the borrowed modal *rado* 'like', which is impersonal.

(91) *rado* [**Ø** *pij-av* *tosara* *jekh* *kuči* *kava*]
like **Ø** drink-1SG morning INDF cup coffee
'I like to have a cup of coffee in the morning.'
(East Slovak, Slovakia, SK-011)

Likewise, Prekmurski dialect of Slovenia has the zero modal complementizer with the borrowed modal *mereš* 'like':

(92) *mereš* [**Ø** *pīj-av* *džēk* *kūči* *kofē* *račas-kro*]
like **Ø** drink-1SG INDF cup coffee morning-GEN
'I like to have a cup of coffee in the morning.'
(Prekmurski, Slovenia, SLO-001)

Borrowed verbs in Romani normally require the use of loan verb adaptation markers. A number of dialects, however, show wholesale borrowing of verbs along with their original (L2) inflectional morphology. Russka Roma, in contact with Russian, and Xoraxani Romani of Bulgaria, in contact with Turkish, are two such dialects. In both of these dialects we find zero modal complementizers with non-adapted complement verbs:

(93) *m-i* *phej* *bašladi* [**Ø** ***bārinmā***] *kana* *thard-e* *amar-e* *khera*
my-F sister began.BOR **Ø** **scream.BOR** when burned-3PL out-PL houses
'My sister began to scream when they burned down our house.'
(Xoraxani, Bulgaria, BG-023)

(94) *o* *la* *nināzdə* *kuveči* [**Ø** ***dönsün*** *khere*]
ART her.OBL was.not.BOR strength **Ø** **walk.BOR** home
'She did not have the strength to walk back home.'
(Xoraxani, Bulgaria, BG-023)

(95) *me* *kam-am* [**Ø** ***ujexatj*** *po* *kurko,* *kaj* *gožo* *i* *kuč*]
I want-1SG **Ø** **leave.BOR** for week where nice and pretty
'I want to go somewhere nice and peaceful for a week.'
(Russka Roma, Russia, RUS-003)

The overall tendencies that condition the absence vs. presence of a complementizer in modal complements across Romani dialects are summarized in Table 4.

Table 4: Zero- vs. overt modal complementizer: tendencies

Ø		***te***
with modals of positive and negative ability	vs.	with other modals
with impersonal modals	vs.	with personal modals
with inherited modals of positive and negative ability	vs.	with borrowed modals of positive and negative ability
with unadapted complement verbs	vs.	with adapted complement verbs
Finnish and Lotfitka Romani	vs.	other dialects

2.5 Combinability issues

Two issues can be flagged in relation to combinability: word order constraints, and the formal combination of complementizer forms. The principal word order constraint applies to the non-factual or modal complementizer *te*, which always appears immediately before the finite verb, thus functioning as an additional, analytical marker of the subjunctive mood:

(96) *me* *kam-om* [*jekvar* *Amerika* ***te*** ***dža-u***]
I want-1SG once America **COMP** **go-1SG.SBJV**
'I want to visit America someday.'
(Čuxny, Estonia, EST-005)

Above we already addressed the distribution of duplex and complex complementizers. These are basically an expansion of the inventory of non-factual complementizer forms, used to capture weak semantic integration and low agent control. Prototypically, they involve a combination of the two main complementizer forms, the factual *KAJ* and the non-factual *te*:

(97) *kam-av* [***kaj*** ***te*** *oddža-ł*]
want-1SG **COMP** **COMP** leave-3SG.SBJV
'I want him to go away.'
(Bergitka, Poland, PL-007)

Recall that *KAJ* represents both the inherited form *kaj* itself, and internal grammaticalization and borrowings that take on the function of the factual complementizer in some of the dialects. Many dialects currently in contact with Russian show the form *sob*, derived from inherited *so* 'what' and the Russian conditional marker *b(y)*, a semi-calque from Russian *čtob(y)*, which is comprised of *čto* 'what' and the conditional/irrealis particle *-b*:

(98) *me* *phend-om* [***sob*** *voj* *jekhatyr* ***te*** *udž-al*]
I told-1SG **COMP** she at.once **COMP** leave.3SG
'I told her to leave at once.'
(Kubanski Servy, Ukraine, UKR-008)

(99) *la-te* *na* *sys* *zor* [***sob*** ***te*** *dža-l* *pale* *khere*]
she-LOC NEG was.3SG strength **COMP** **COMP** go-3SG back home
'She did not have the strength to go back home.'
(Lithuanian Romani, LT-005)

Borrowed prepositions with the final meaning 'for' or 'in order to' may also combine with non-factual *te* in cases of predication pairs that are on the weak end of the semantic integration continuum. In the following cases, the combinations *ja ti* in a Greek Romani dialect and *za te* or *za da te* in Bulgarian Romani dialects replicate the respective constructions *ja na* and *za da* in Greek and Bulgarian:

(100) *i* *dar* *ker-ela* [***ja*** ***ti*** *borin-el* *ti* *rov-el*]
ART fear did-3SG **COMP** **COMP** can-3SG COMP cry-3SG
'The fear made her cry.'
(Sofades, Greece, GR-004)

(101) *voj* *pučlj-as* *ma* *so* *te* *ker-el* [***za*** ***da*** ***te***
she asked-3SG me.OBL what COMP do-3SG **COMP** **COMP** **COMP**
ker-el *pobut* *love*]
make-3SG more money
'She asked me what to do to earn some more money.'
(Kalajdži, Bulgaria, BG-007)

Note that the word order restriction mentioned at the beginning of this section, namely that of non-factual *te* immediately preceding the verb, determines the order of complementizers in a duplex form: the factual *KAJ*-type complementizer (including *kaj* and the various loans and calques) always precedes the non-factual *te*.

2.6 Non-complementizing functions of complementizer forms

Both 'prototype' complementizers, factual *kaj* and non-factual *te*, are aligned with other clause-combining and modality functions, respectively. Factual *kaj* serves in interrogative clauses as an interrogative pronoun 'where' and as such it can introduce embeddings that describe location:

(102) *me ne-bi puč-av tut te džan-av kaj odova*
I NEG-COND ask-1SG you.OBL COMP know-1SG where this
'I wouldn't ask you if I knew where it is.'
(Kosovan Romani, Serbia, YU-018)

It is also the most common relative pronoun in Romani dialects:

(103) *dikhlj-om o kher kaj bori-es-as andar les-te*
saw-1SG ART house which talk-2SG-REM about him-LOC
'I saw the house that you were talking about.'
(Ursari, Romania, RO-004)

Non-factual *te* has a general modality function that can mark the optative/imperative, especially in interrogative clauses:

(104) *So te ker-av?*
what COMP do-1SG
'What shall I do?'
(general)

In clause combinations, it can introduce dependent aspectual constructions, and is also the most common inherited conditional conjunction in Romani:

(105) *palo pančminutora počnisard-a te čero-l lafi*
after five minutes started-3SG COMP make-3SG words
'After five minutes he started to talk.'
(Gurbet, Macedonia, MK-001)

(106) *te av-en man lōve, tu-ke d-os le*
COMP come-3PL me.OBL money you-DAT give-1SG.REM them.OBL
'If I had some money I would give it to you.'
(Gurvari, Hungary, HU-007)

(107) *ta v-es dikk-a-t*
COMP come-2SG see-1SG-you
'If you come, I shall see you.'
(Molise Romani, Italy, IT-010)

Finally, in combination with the preposition *bi* 'without', non-factual *te* also introduces adverbial clauses that express negative circumstance:

(108) *sar* *možі* *dž-as* *dži* *ando* *foros* *bi* *te* *molisar-av* *məndr-ə*
how can go-1PL until in town without COMP ask-1SG my-OBL
phral-es *e* *kolake*?
brother-OBL ART car
'How can we get to town without asking for my brother's car?'
(Kalajdži, Bulgaria, BG-007)

2.7 Diachrony

The etymology of *kaj* is fairly straightforward: it is the inherited Romani interrogative 'where'. Its use as a relativizer resembles the grammaticalization path of similar forms in other contact languages of the Balkans, most notably Greek, as does its extension to factual complements. The etymology of *te* is less obvious. It is not a cognate of Domari *ta* 'in order to', which is borrowed from Arabic. There are however other languages in the Near East that employ *ta* in purpose clauses, e.g. Kurdish and Neo-Aramaic, where it appears to originate in the Iranian preposition *ta* 'until'. A deictic etymology for *te* has been considered by various authors, linking it with the Old Indo-Aryan pronoun *ta-* (Pobożniak 1964: 58), the Hindi correlative *to* (Pott 1845: 281), or Old Indo-Aryan *iti* 'so' (Sampson 1968 [1926]: 363). The correlative function is an attractive etymology as it can be related to the semantic dependency that characterises Romani *te* (see Matras 1994: 231–233, Matras 2002: 180).

The factual complementizer *kaj* is prone to borrowing. Matras (2002: 179–80) lists three main geographic zones where *kaj* has been replaced. The first zone on the list comprises Vlax dialects of Romania, Moldova and elsewhere, where *kaj* has been replaced by the Romanian complementizer *ke*. This change also affects the Balkan type Ursari dialect of Romania. The second zone comprises Arli and Southern Vlax varieties of Greece, where *kaj* has been replaced by the Greek *oti*. The change also affects the Dendropotamos and Agia Varvara Vlax varieties, which are spoken by more recent immigrants to Greece; this illustrates the high susceptibility to contact-induced change of the domain of factual complementation. The third zone comprises the Central dialects of Slovakia and Hungary, where *kaj* is in the process of being replaced by the Hungarian loan *hod*/*hodž*/*hod'*/*hot*/*hoj*. Consider the following examples from these three zones.

(109) *voj* *phendj-as* [***ke*** *či* *prinžan-el* *khanikas*]
she said-3SG **COMP** NEG know-3SG nobody
'She said that she did not know anyone.'
(Kaldaraš, Romania, RO-008)

(110) *on* *patjan-ile* [***ke*** *som* *ko* *birtos*]
they thought-3PL **COMP** am at bar
'They thought that I was in the pub.'
(Ursari, Romania, RO-004)

(111) *asund-om* [***oti*** *bes-ena* *javer* *da* *roma* *ate*]
heard-1SG **COMP** live-3PL other also Roma here
'I heard that other Roma live here as well.'
(Sofades, Greece, GR-004)

(112) *iērzindj-am* [***hodj*** *valesosko* *bajo* *hi* *la*]
felt-1PL **COMP** something bad is her.OBL
'We felt that something was wrong with her.'
(Gurvari, Hungary, HU-007)

(113) *džan-el* *pes* *pal* *les-te* [***hoj*** *hin-o* *but* *barval-o*]
know-3SG REFL about him-LOC **COMP** is-M very rich-M
'It is known that he is very rich'
(East Slovak, SK-011)

Some additional dialects also have a borrowed factual complementizer. The Lombard Sinti dialect of Italy uses Italian *ke*, many Bulgarian dialects have borrowed Turkish *ani*, and the varieties spoken in Finland use either Finnish *et(tä)* or Swedish *at*:

(114) *sperar-ava* [***ke*** *sigo* *finar-ela* *ta* *d-el*]
hope-1SG **COMP** soon stop-3SG COMP give-3SG
'I hope it stops raining soon.'
(Lombard Sinti, Italy, IT-011)

(115) *ašund-em* [***ani*** *si* *tumen* *šukar* *buki*]
heard-1SG **COMP** is you.PL.OBL nice job
'I heard that you have a good job.'
(Kalburdzhu, Bulgaria, BG-008)

(116) *tenkav-aa* [***at*** *me* *vojpuv-aa* *presav-es* *tela* *louve* *tu-kke*]
think-1SG **COMP** I can-1SG pay-2SG back money you-DAT
'I think that I will be able to pay you back.'
(East Finnish Romani, FIN-008)

A slightly more complex replacement of the inherited factual complementizer is found in many dialects that are in contact with Russian and Ukrainian. These include many, but not all, of the varieties of the Northeastern Romani group, namely Russka Roma, Lithuanian Roma and Estonian Roma, as well as some of the Servi and Vlaxurja type dialects of Ukraine. In these dialects *kaj* has been

completely replaced by *so* 'what', following the Russian model, where the factual complementizer *čto* has formal syncretism with the interrogative *čto* 'what':

(117) *me dužakir-av* [***so*** *mir-o čavo jav-ela s minuty na minute*]
I wait-1SG **COMP** my-M son come-3SG from minute to minute
'I expect my son to come here any minute.'
(Russka Roma, Russia, RUS-003)

(118) *jov sys dasav-o sasto* [***so*** *vasten-sa vraskir-l-as sastruno*]
he was so-M strong-M **COMP** hands-INS bend-3SG-REM iron
'He was so strong that he could bend the iron with his hands.'
(Servi, Ukraine, UKR-003)

The Northeastern dialects of Poland (Polska Roma, Bergitka, Polish Xaladytka) do not have *so*, but rather use the inherited *kaj*, which is not surprising, because Polish, unlike Russian, does not have syncretism of the interrogative 'what' (Polish *co*) and the factual complementizer (Polish *że*). This contact-related difference between dialects under Russian vs. Polish influence has been discussed in Tenser (2008: 206). It should be noted, that the Polish Xaladytka dialect is spoken by a group of relatively recent migrants from the Russian-speaking territories, and has most likely used the Russian model *so* until about two generations ago. This quick change from *so* to *kaj* in Polska Xaladytka was due to inter-dialectal contact and once again illustrates the high volatility of non-modal complementation.

Lotfitka Romani (Latvia) also shows an innovation here; it has replaced *kaj* with *sy* 'how'. This seems to be a calque from Latvian, stemming from the non-differentiation of Latvian interrogative *kā* 'how' and factual complementizer *ka*:

(119) *me žakir-a* [***sy*** *m-o čavo vraši te ja-l paše*]
I wait-1SG **COMP** my-M son soon COMP come-3SG back
'I expect my son to come back any minute.'
(Lotfitka, Latvia, LV-005)

In several samples two different forms of the factual complementizer coexist. In one sample from Bulgaria (Xoraxani) there is an alternation between *či* (of Bulgarian origin) and *ani* (of Turkish origin), and in various samples from Romania inherited *kaj* coexists with the borrowed *ke*; one of the Ukrainian Servi samples has both *kaj* and *so*. Mostly in such cases there does not seem to be a clear-cut grammatical or semantic distribution of the two forms. Thus in the same sample we find:

(120) *džan-ava* [***ani*** *bu iki senede sja tut but xəzmeči*]
know-1SG **COMP** last two years was you.OBL much work
'I know that you had a lot of work during the past two years.'
(Xoraxani, Bulgaria, BG-015)

(121) *džan-ava* [***či*** *odva* *vakerd-a* *tumen-ge* *kal-es*]
know-1SG **COMP** he said-3SG you.PL-DAT this-OBL
'I know that he said this to you (pl).'
(Xoraxani, Bulgaria, BG-015)

In the Servi sample where *kaj* and *so* co-exist, there seems to be a tendency to use *so* with complements that have a stronger independent truth value, and *kaj* with those that have a weaker independent truth value:

(122) *man-ge* *na* *d-el-pe* [***kaj*** *varokon* *adava* *skjerd-a*]
me-DAT NEG give-3SG-REFL **COMP** someone this did-3SG
'It does not seems to me that anyone did it.'
(Servi, Ukraine, UKR-003)

(123) *me* *phendj-om* *les-ke* [***so*** *banza* *čekir-el-pe* *de* *deš* *štunde*]
I told-1SG him-DAT **COMP** store open-3SG-REFL at ten hours
'I said to him that the shop opens at 10.'
(Servi, Ukraine, UKR-003)

3 Summary

Romani relies on combining finite clauses in complex predications, and complementizers play the key role in identifying the nature of the semantic links between main and embedded (complement) clauses. The choice of complementizer reflects the degree of semantic integration between the two clauses. The principal distinction is one between factual complementation, and non-factual or modal complementation. In following this basic typology of complement clauses, represented primarily by the choice of complementizer and correlating features such as tense-mood selection and word order, Romani aligns itself with the linguistic area of the Balkans, where most of the languages show a similar distinction. Historically, this goes back to the formation of Early Romani as a language of Early New Indo-Aryan heritage that appears to have undergone significant typological shift in contact with the languages of the Balkans, especially Greek, after the settlement of Romani-speaking populations in the Byzantine Empire in the period around the tenth century CE (cf. Matras 1994, 2002 and Elšík & Matras 2006).

The factuality distinction manifests itself in Romani prototypically through the choice between the complementizers *kaj* (factual) and *te* (non-factual, modal). Factual complements accompany verbs of perception and utterance, and show independent selection of tense-mood and word order patterns. Non-factual, modal complements accompany verbs of intent, command, and manipulation, with purpose clauses aligning themselves with the same type. They normally

appear in the subjunctive (though some dialects show other patterns), and the complementizer *te* that introduces them normally appears in the position immediately preceding the finite verb of the complement clause. Clause integration in complementation is arranged on a continuum of semantic integration, which is relevant primarily for non-factual complements. These can range from immediate, direct or tightly integrated modal complements, to those that are less tightly integrated. The criteria for semantic integration include agent control over the target action (the action depicted by the embedded, complement predicate), control over the manipulee (in different-subject constructions), the degree of event independence (especially in purpose clauses) and the likelihood that the target predication can be accomplished. Dialects that show sensitivity to this semantic continuum of non-factual complements in the form of a variety of complementizers tend to show a correlation between the structural complexity of the complementizer itself and the tightness of semantic integration; that is, tight integration is more likely to be represented by a simplex complementizer, while weaker semantic integration is flagged by a duplex or even triplex complementizer, in which the default non-factual complementizer *te* is extended by a 'reinforcer' form (often the factual complementizer, or a preposition, or both). In terms of diachrony, the inherited pattern emerging from Early Romani and which is continued in some form or other in most dialects shows a grammaticalization path that derives complementizers ultimately from deictic forms. More specifically, the non-factual *te* has its roots in all likelihood in an ancient correlative. The factual complementizer *kaj* derives from an interrogative turned relativizer. Contact developments in the various dialects of Romani attest to a high susceptibility of this factual complementizer to borrowing, and in many dialects it is directly replaced by borrowed complementizers, while the inherited non-factual complementizer generally remains stable.

Country abbreviations in RMS sample codes

AL	Albania
BG	Bulgaria
CZ	Czech Republic
EST	Estonia
FIN	Finland
GR	Greece
HR	Croatia
HU	Hungary
IT	Italy
LT	Lithuania
LV	Latvia
MD	Moldova
MK	Macedonia
MX	Mexico
PL	Poland
RO	Romania
RUS	Russia
SK	Slovakia
UKR	Ukraine
YU	Yugoslavia

References

Boretzky, Norbert. 1996. The 'new infinitive' in Romani. *Journal of the Gypsy Lore Society*, Fifth Series, 6. 1–51.

Boye, Kasper. 2012. *Epistemic meaning. A crosslinguistic and functional-cognitive study*. Berlin: De Gruyter Mouton.

Cristofaro, Sonia. 2003. *Subordination*. New York: Oxford University Press.

Elšík, Viktor & Yaron Matras. 2006. *Markedness and language change. The Romani sample*. Berlin: Mouton de Gruyter.

Elšík, Viktor & Yaron Matras. 2009. Modality in Romani. In Björn Hansen & Ferdinand de Haan (eds.), *Modals in the languages of Europe*, 267–324. Berlin: Mouton de Gruyter.

Givón, Talmy. 1990. *Syntax. A functional-typological introduction*, vol. II. Amsterdam: John Benjamins.

Matras, Yaron. 1994. *Untersuchungen zu Grammatik und Diskurs des Romanes. Dialekt der Kelderaša/Lovara*. Wiesbaden: Harrassowitz.

Matras, Yaron. 1999. The speech of the Polska Roma: some highlighted features and their implications for Romani dialectology. *Journal of the Gypsy Lore Society,* Fifth Series, 9(1). 1–28.

Matras, Yaron. 2002. *Romani: A linguistic introduction*. Cambridge: Cambridge University Press.

Matras, Yaron 2004. Typology, dialectology and the structure of complementation in Romani. In Bernd Kortmann (ed.), *Dialectology meets typology*, 277–304. Berlin: Mouton de Gruyter.

Matras, Yaron. 2005. The classification of Romani dialects: A geographic-historical perspective. In Barbara Schrammel, Gerd Ambrosch & Dieter W. Halwachs (eds.), *General and applied Romani linguistics*, 7–26. Munich: Lincom Europa.

Pobożniak, Tadeusz. 1964. *Grammar of the Lovari dialect*. Kraków: Państwowe wydawnictwo naukowe.

Pott, August. 1844–1845. *Die Zigeuner in Europa und Asien. Ethnographisch-linguistische Untersuchung vornehmlich ihrer Herkunft und Sprache*. Halle: Heynemann.

Sampson, John. 1968 [1926]. *The dialect of the Gypsies of Wales, being the older form of British Romani preserved in the speech of the clan of Abram Wood*. Oxford: Clarendon Press.

Tenser, Anton. 2008. *Northeastern group of Romani dialects*. Manchester: University of Manchester dissertation.

Part II: **Non-Indo-European languages**

Xabier Artiagoitia and Arantzazu Elordieta

On the semantic function and selection of Basque finite complementizers

1 Introduction

Basque finite complementizers can be described as suffixes on finite verb forms. In this paper, we hope to make clear that Basque complementizers have semantic values associated with a declarative vs. question contrast, factivity, polarity, and mood or propositional content evaluation. We hold that these values can be expressed as morphosyntactic features on complementizers in Basque, and that some of these values are expressed by means of the subjunctive mood in Romance.

We have organized our article as follows: we start by providing a brief introduction to the main linguistic and sociolinguistic features of the language; the second part is a more thorough description of the semantic and syntactic functions of the different Basque finite complementizers, some of which are restricted to certain dialect(s); Section 3 sketches a feature value analysis from a generative perspective of the two most common and productive complementizers (viz. *-ela* and *-en*); Section 4 tackles two more intricate issues: (a) the semantic import of the Western Basque complementizer *-enik*; and (b) the surprising alternation between *-ela* and *-en* in lexically selected subjunctive complements. Our conclusions are presented in Section 5.

1.1 Some phylogenetic and sociolinguistic aspects of the Basque language

From the evidence gathered in the literature so far, it is fair to assert that Basque is genetically unrelated to any other living language. In the literature there have been different proposals about possible origins and connections, but they all have been disregarded for incontestable reasons, both empirically and theoretically motivated (see Trask 1995, also Hualde 2003 for a brief summary). The only proposal that has borne fruitful results is the one which connects Basque with Aquitanian (see Mitxelena [1954] 1985; Gorrotxategi 1984, 1995), a dead language spoken in the southwestern part of Gaul during the Roman period. There are no attested long texts in Aquitanian, but a few hundred personal names and names of divinities, embedded in Latin texts. The phonology and morphology of these

names share a striking similarity with Basque words (some of them are identical to the forms independently reconstructed for Proto-Basque), hence most Basque scholars have concluded that Aquitanian is the closest ancestor of Basque. Apart from this ancient relationship, Basque is the only non-Indo-European language in Western Europe.

Today there are about 700,000 fluent speakers of Basque, divided in three administrative territories across two countries (cf. Eusko Jaurlaritza 2009, Nafarroako Gobernua 2009, Baxok 2007: 1) the Autonomous Community of the Basque Country in Spain, 2) the Community of Navarre in Spain, and 3) across the French border, the Basque speaking area integrated by the regions of Lapurdi (French Labourd), Low Navarre (Basse Navarre) and Zuberoa (French Pays de Soule), which together comprise the western half of the Département des Pyrénées-Atlantiques. Almost without exception, all Basque speakers are completely bilingual in Spanish or French, and the number of speakers is gradually increasing among the younger generations in the Autonomous Community of the Basque Country and in Navarre as a result of the use of Basque in the educational system (Hualde 2003: 3). Unfortunately, the situation at the other side of the Pyrenees is less optimistic; Basque used to be the majority language three or four centuries ago, but today the language is under attrition, probably induced by the lack of official recognition (see Oyharçabal 1997). Quite on the contrary, in the Spanish Basque speaking areas Basque has official status (together with Spanish) in the Autonomous Community of the Basque Country, which is conformed by the provinces of Bizkaia, Gipuzkoa and Araba. In the Community of Navarre it also has official status, but it is more limited.[1]

1.2 The standard variety of the language

The standardization of the language is relatively recent (a bit more than four decades long), and although its successful implementation in the educational system and in the media has brought with it a great uniformity in the use of written Basque, there are still noticeable differences among Basque speakers of different dialects in their pronunciation (phonology), in some aspects of the morphology, in their lexicon and, as we will show here in relation to the complementizer system, also in their syntax. As Hualde (2003: 3) notes, this is expected for a language that is spoken in a mountainous area and which until very recently

1 This is due to a further subdivision the Government of Navarre makes on distinct areas depending on the use of Basque in such areas. Accordingly, Basque is officially recognized only in those parts of the territory which the Government of Navarre considers to be mostly Basque speaking.

has lacked both official status and a standard form. With respect to the classification of Basque dialects, we will adopt Zuazo's (1998, 2008) more recent proposal, according to which there are six dialects, namely: Western (= Biscayan in other classifications), Central (= Gipuzkoan in other classifications), Navarrese, Lapurdian-Navarrese, Zuberoan and, finally, Eastern Navarrese, a now extinct dialect. As pointed out in the thorough descriptive grammar of Basque cited above (Hualde & Ortiz de Urbina 2003: 3), many important isoglosses coincide at present roughly with the Spanish-French border. Hence it is of relatively common usage to refer to dialects spoken in France (i.e. Lapurdian-Navarrese and Zuberoan) as *northern*, as opposed to the *southern* ones spoken in Spain (i.e. Western, Central and Navarrese). In addition, in some instances certain linguistic features are shared by northern dialects and Navarrese, in opposition to the other dialects. We will see an example of this when we discuss the dubitative complementizer *-en* 'that' in Section 2.2. For these cases, we will use the term *eastern* to refer to the former group of dialects (i.e. northern (Lapurdian-Navarrese and Zuberoan) and Navarrese), and *western* (in lower case) to refer to the latter dialects (i.e. Western and Central dialects).

We hope to have made clear the terminology we will follow in the article to refer to a specific distinctive feature which sets one particular dialect or a set of dialects apart from the others. This aside, the majority of Basque examples used in the present article are given in standard Basque, unless stated otherwise.

1.3 Basic linguistic features of Basque

In this section we will mention some general properties of Basque we consider relevant and which will facilitate the following discussion of the data to those readers non-familiar with the language.

1.3.1 Ergativity and agreement

Basque is an ergative language. This means that it makes a distinction between subjects of transitive predicates and subjects of intransitive predicates, both in the nominal and in the verbal domain. In the case system, there are three cases that instantiate the grammatical relations established among the arguments of a predicate, namely, the ergative *-k*, borne by agentive subjects, (1a) and (1b), the

absolutive *-Ø,* borne by objects and subjects of intransitive predicates, (1a) and (1c), and the dative *-(r)i*, associated with goal arguments, (1b):[2]

(1) a. *Jon-ek* *kafe-a-Ø* *edan* *du*
Jon-ERG coffee-ART-ABS drink AUX
'John has drunk coffee'
b. *Jon-ek* *lagun-a-ri* *esan* *dio*
Jon-ERG friend-ART-DAT say AUX
'John said (it) to his friend'
c. *Jon-Ø* *gaur* *heldu* *da*
Jon-ABS today arrive AUX
'John has arrived today'

Likewise, finite verbal forms carry person agreement features associated with up to three arguments, besides being marked for tense and mood. This holds both in the case of periphrastic forms, consisting in a non-finite verb form followed by a finite auxiliary – most of the examples provided in the text –, as well as in the so-called synthetic verb forms, a handful of verbs in which the lexical verb and the inflectional morphology merge together – some illustrative examples are given in (7b) and (8c) below –.[3] What is remarkable is the fact that agreement morphemes corresponding to subjects marked with ergative are distinct from agreement markers corresponding to subjects bearing absolutive, a result which squares well with the ergative pattern we observed in the case system in (1).[4] As an illustration, consider the sentences in (2). The relevant comparison is the following, highlighted in bold face in the examples: compare the agreement marker (*-t*) corresponding to the 1SG subject (*ni-k* 'I') of the transitive verb *ekarri* 'bring' with the agreement marker (*n-*) corresponding to the 1SG subject (*ni-Ø* 'I') of the unaccusative verb *etorri* 'come' in (2b); the very same marker is used for object agreement with the 1SG (*ni-Ø*) in transitive clauses in 2c. In addition, (2d) illustrates that the finite verb also agrees with the dative indirect object:

2 For other proposals which defend the view that in the western Basque dialects the case-marking pattern is semantically based on notions related to agenthood, see Aldai (2008, 2009) (and, from a different perspective, also Laka 2005, 2006). The main arguments of these proposals lie on the dialectal differences existing on the eastern-western axis regarding the subject case-marking of lexically-simple unergative verbs (i.e. whether it bears ergative or absolutive case). A detailed discussion of these facts is beyond the scope of this article, hence we refer the reader to the cited references for further details.

3 Although in earlier stages of the language the situation was different, in modern Basque verbal morphology is overwhelmingly periphrastic.

4 As is illustrated in the examples in (2b–c), in present tenses the absolutive agreement marker precedes the tense marker and the root. However, when the phrase bearing absolutive refers to 3.p. (as in the examples [2a] and [2d]), no overt agreement morpheme surfaces and instead the tense marker *d(a)-* occurs as the first affix preceding the root (see Laka 1993 for details).

(2) a. *Ni-k laguna-Ø kotxe-z ekarri d-u-**t***
I-ERG friend-ABS car-by bring PRS-ROOT-1SG.ERG
'I have brought my friend by car'
b. *Ni-Ø kotxe-z etorri **n**-a-iz*
I-ABS car-by come 1SG.ABS-PRS-ROOT
'I have come by car'
c. *Jon-ek ni-Ø kotxe-z ekarri **n**-a-u-Ø*
Jon-ERG I-ABS car-by bring 1SG.ABS-PRS-R-3SG.ERG
'John has brought me by car'
d. *Ni-k lagun-a-ri liburu bat-Ø eman d-i-**o**-**t***
I-ERG friend-ART-DAT book one-ABS give PRS-ROOT-3SG.DAT-1SG.ERG
'I have given a book to my friend'

Thus, these data do not only constitute an illustration of the three-way verbal agreement Basque reveals, but also show that Basque displays a fully ergative morphology in its grammatical system.

1.3.2 A pro-drop language

A further general property of Basque which deserves mentioning is the fact that Basque is an extended *pro-drop* language, that is, it allows not only null subjects, but, also null objects. Accordingly, the three verbal arguments in subject, direct object and indirect object positions can be omitted (cf. Ortiz de Urbina 1989a):

(3) *Esan d-i-zu-t*
say PRS-R-2SG.DAT-1SG.ERG
'(I) have said (it) (to you)'

In fact, argument drop is also found in non-finite clauses, as illustrated in (4):

(4) *Nahi d-u-zu* [*Jon-ek/e zu/e etxe-ra erama-te-a*]*?*
want PRS-ROOT-2SG.ERG Jon-ERG/e you/e home-ALL bring-NMLZ-ART
'Do you want John/(me)/(him) to bring (you)/(him) home?'

This suggests that the possibility of having null arguments is not directly dependent on the existence of rich verbal agreement (cf. Elordieta 2002; Duguine 2012 for some proposals in this direction). Furthermore, the embedded non-finite complement clause in (4) shows that null arguments (indicated by *e*) can freely alternate with lexical DPs (i.e. the referential expression *Jon* and the pronoun *zu* 'you'), which seems to suggest that morphological agreement and Case assignment are distinct issues (Elordieta 2001: 36–38).

A further linguistic feature relevant in Basque is constituent order, given that Basque is assumed to be head-final, but yet allows considerable variation among

the ordering of sentential constituents. A brief and very general description of the facts is provided in the next subsection.

1.3.3 Word order

We adopt the standard assumption that S (IO) O V is the underlying base constituent order in Basque, which corresponds to the unmarked or neutral order, where no sentential element is pragmatically salient in terms of new/known information (see de Rijk 1969; Goenaga 1980; Euskaltzaindia 1985; Ortiz de Urbina 1989a; Elordieta 2001, among many others). Ever since Greenberg (1963), Basque has been characterized as a language with dominant, but not exclusive, SOV word order. Indeed, Basque conforms to some of the generalizations about SOV languages which have been stated in the typological literature: for instance, the fact that Basque has postpositions, (5a) and (5b), that the auxiliary follows the main verb, (5a), (5b) and (5c) (cf. de Rijk 1969), and that the complementizer follows the subordinate clause, both in complement (5b) and relative clauses (5c):

(5) a. *Nire* *lagun-a* *eskola-ra* *joan* *da*
My friend-ART school-ALL go AUX
'My friend has gone to school'

b. [*Nire* *lagun-a* *eskola-ra* *joan* *d-ela*] *esan* *didate*
my friend-ART school-ALL go AUX-COMP say AUX
'They have told me that my friend has gone to school'

c. *Jonek*$_i$ [e_i *erosi* *du-en*] *kotxe-a* *erakutsi* *digu*
Jon-ERG buy AUX-COMP car-ART show AUX
'John has shown us the car that he has bought'

Basque also displays a head-final pattern within the nominal phrase; thus, the article is attached to the rightmost element of the noun phrase, (5a), (5b) and (5c), and relative clauses must precede their nominal heads, (5c) (cf. de Rijk 1969). Nonetheless, in addition to this, there is relatively free variation in the surface order of the constituents of a declarative sentence in Basque. This variation relies heavily on the focus interpretation of the sentence; that is to say, the constituent in focus will appear preceding the finite verb, and very often all other phrases will occur postverbally, although they may occur preverbally as topics (Ortiz de Urbina 1989a; Elordieta 2001; Hualde & Ortiz de Urbina 2003, among others). In a sense, Basque displays the behavior of a "residual" verb-second language in focus constructions, similar to the V2 phenomenon of many Germanic languages (cf. Ortiz de Urbina 1989a, 1995), although the comparison does not hold in general (Basque is not strictly V2 in declarative non-focus constructions). It holds as for the idea that in focus constructions the focused phrase and, subsequently,

the finite verb are fronted to a sentence-initial position (be it CP, FocP or a similar position), yielding a V2 order. (5d) exemplifies a construction where the subject is focused. In (5d) the sequence is S V Aux PP, instead of the neutral or unmarked S PP V Aux in (5a). Fronting a focused constituent as well as the finite V is a strategy Basque employs in order to focalize a constituent:

(5) d. *Nire* *lagun-a* *joan* *da* *eskola-ra*
My friend-ART go AUX school-ALL
'My FRIEND has gone to school'

Finally, as we will see in Section 2.1, negation also prompts auxiliary inversion effects in the order of the sentence, some of which will be discussed in relation to complementizers (see [11] below) .

Therefore, following the tradition, we will assume that Basque has S O V Aux as its basic word order, which is to be identified with the order found in unmarked (pragmatically neutral) contexts (see references cited at the beginning of this section).[5]

After having presented a very basic linguistic characterization of Basque, next we will proceed to focus on the kernel of the matter, namely the description of the system of finite complementizers in Basque.

2 A snapshot of the Basque complementizer system

The analysis that we will be presenting in the following sections focuses on finite complementizers, but we consider that a brief mention of other major complementation strategies employed in the language, more specifically, of non-finite complement clauses, will be of interest to the reader.[6] Basically, emotive verbs (and emotive predicates in general), perception verbs and the command-type of verbs select for non-finite complement clauses, which surface as nominalizations, (6a), or subordinate clauses headed by the affix *-tzen*, (6b) or *-tzeko*, (6c) (see Goenaga 1984, 1985 for an analysis of Basque non-finite complementation, Ortiz de Urbina 1989a for nominalizations, also Arteatx 2011 for a different account on the syntactic structure of complements of perception verbs). It is true,

5 This seems to be a reliable criterion to identify the basic word order in Basque. For a discussion of this and other additional criteria, see, for instance, Croft (1990), Dryer (2007).

6 We thank the editors of this volume for their recommendations on this issue.

however, that *-tzeko* complements to directive verbs can also take the shape of a finite subjunctive complement (see [37b] in Section 4.2):

(6) a. [*Gerrako filmak ikus-tea*] *gustatzen zait*
war films watch-NMLZ like AUX
'I like to watch war films'

b. [*Baker-en azken liburu-a irakur-tzen*] *ikusi zaitut*
Baker-GEN last book-ART read-TEN see AUX
'I've seen you reading Baker's most recent book'

c. *ama-k* [*garaiz hel-tzeko*] *agindu dit*
mother-ERG on time arrive-TZEKO order AUX
'Mother has requested (from) me to arrive on time'

For reasons of limited space, we cannot provide a full picture of the non-finite complementation system, but for the purposes of this chapter it may be relevant to highlight that, as is often noted in the literature, the semantic type of non-finite complements correlates with state-of-affairs, whereas finite complementation correlates with propositions. Indeed, the propositional content of the complement clauses headed by the finite complementizers we will introduce next will be at stake when we deal with the different interpretations carried out by those complement clauses in relation to their truth value and factivity.

In what follows we will make a division between the finite complementizers *-en* and *-ela*, which are common to all dialects and will be described in Section 2.1, and the rest, which show a different range of dialectal variation and will be described in Section 2.2.

2.1 Common complementizer system (all dialects)

Both *-ela* and *-en* are used in all dialects of Basque. The citation form of the two suffixal complementizers is subject to some variation, given that they may both surface as just *-la* or *-n* depending on the shape of the last morpheme of the verb. Sometimes the two morphemes are referred to as just *-la* or *-n*, or even *-(e)la* and *-(e)n* (implying that the initial vowel is some kind of epenthesis, contrary to truth). To make a long story short, the *-la* an *-n* allomorphs show up after: (a) any person marker but the third person singular (i.e. Ø); (b) a stem final *-a* that does not belong to the copula or auxiliary verb *izan* 'be'; (c) the past tense suffix *-n*. The allomorphs *-ela* and *-en* are used elsewhere. A few examples will illustrate the distribution:

(7) *-la* and *-n* allomorphs

a. ekarri {du-t, du-gu} + {-la, -n} → ekarri {*duda-la, dugu-la; duda-n, dugu-n*}
bring {have-1SG, have-1PPL}
'{I, we} have brought it' '{that, whether} {I, we} have brought it'

b. darama-Ø + {-la, -n} → *darama-la, darama-n*
carry-3SG
'(s)he carries' 'that (s)he carries it, whether (s)he carries it'

c. gustatu zitzaida-n + {-la, -n} → gustatu *zitzaida-la, zitzaida-n*
like AUX-PST
'It pleased me' that it pleased me, whether it pleased me'

(8) *-ela* and *-en* allomorphs

a. ekarri du-Ø + {-ela, -en} → ekarri *du-ela, du-en*
bring have-3SG
'(s)'he has brought it' 'that (s)he has brought it, if (s)he has brought it'

b. etorri {da, zara}+ {-ela, -en} → etorri *{d-ela, zar-ela}*, etorri *{zar-en, d-en}*
come AUX(3SG), AUX(2SG)
'{he, you} have come' 'that, whether {he, you} have come'

c. dago + {-ela, -en} → *dago-ela, dago-en*
stay(3SG)
'(s)he stays' 'that (s)he stays, whether (s)he stays'

In other words, the past tense marker *-n* is elided before *-la* and a single *-n* is found in the case of a combination [past morpheme + complementizer] in (6c). With respect to the 1SG marker *-t*, it is agreed that the old and underlying form is **-da*, which becomes *-t* (<*d*) after a vowel truncation rule, if no other suffix is present (e.g., in main clauses).

Roughly speaking, *-ela* is basically a declarative complementizer, and *-en* can be described as a wh-complementizer, present in indirect questions with or without interrogative pronouns. None of them can ever be deleted, but they differ with regard to the possibility of triggering auxiliary inversion when negation is present. *-ela* induces Aux-inversion, whereas *-en* allows both inverted and non-inverted orders. A summary chart is given in Table 1:[7]

7 The % symbol in Table 1 is meant to refer to speaker variation, rather than true optionality.

Table 1

Complementizer	Aux-inversion with negation	Semantic Clues	In non-complement clauses	Deletable
-ela	yes	declarative	yes	no
-en	%	indirect interrogatives and exclamatives	yes	no

The examples in (9) show, for a couple of representative verbs like *sinetsi* 'believe', *zalantza eduki* 'doubt', *galdetu* 'wonder', that the two complementizers are for the most part in complementary distribution (unless the matrix verb allows for both declarative and interrogative complements, as is the case of *esan* 'say' or *erantzun* 'answer', cf. [9c]):

(9) a. *sinesten dut* [*Jon etorri* {*d-ela* /**d-en*}]
believe AUX Jon come AUX-COMP AUX-COMP
'I believe {that, *whether} John came'

b. {*zalantza daukat, galdetu dut* [*Jon etorri* {**d-ela* /*d-en*}]
doubt have.1SG ask AUX Jon come AUX-COMP AUX-COMP
'I {doubt, asked} {*that, whether} John came'

c. [*Jon etorri* {*d-ela* /*d-en*}] {*esan, erantzun*} *nezake*
Jon come AUX-COMP AUX-COMP say, answer AUX
'I could say {that , whether} John came'

d. [*Jon nolako ergela d-en*] *kontatu dit*
Jon how idiot AUX-COMP report AUX
'He reported {how stupid John is, to which degree John is stupid}'

Example (9d) illustrates that a wh-complement headed by *-en* and an overt interrogative pronoun can also have an exclamatory interpretation.

The ungrammaticality of (10a) and (10b) below shows that in no case is either the complementizer *-ela* or *-en* deletable:

(10) a. **sinesten dut* [*Jon etorri da*]
believe AUX Jon come AUX
'I believe John came'

b. *{*zalantza daukat, galdetu dut*} [*Jon etorri da*]
doubt have ask AUX Jon come AUX
'I {doubt, asked} John came'

One interesting syntactic difference between the two complementizers has to do with word order: like in matrix negative sentences, *-ela* complements trigger Auxiliary Inversion, whereas *-en* complements only do so optionally (see Ortiz de Urbina 1989b; Laka 1991). The example (11a) shows that negation triggers Aux-Inversion in a matrix clause; example (11b) illustrates that no Auxiliary Inversion

is required with *-en*, and (11c) indicates that *-ela* (and, optionally, *-en*) triggers Aux-Inversion:

(11) *matrix negation*:
a. [*Jon etorri da*] *vs.* [*Jon ez da etorri*] *but* *[*Jon etorri ez da*]
Jon come AUX Jon NEG AUX come Jon come NEG AUX
'John came' 'John did not come'
complement clause without Aux-Inversion:
b. [*Jon etorri ez* {*??d-ela /d-en*}] *esan dut*[8]
Jon come NEG AUX-COMP AUX-COMP say AUX
'I said {that /whether} John did not come'
complement clause with Aux-Inversion:
c. [*Jon ez* {*d-ela /d-en*} *etorri*] *esan dut*
Jon NEG AUX-COMP AUX-COMP come say AUX
'I said {that /whether} John did not come'

Outside complement clauses, *-en* is used mainly as a relativizer in the formation of prenominal relative clauses, as in (12a) below, whereas *-ela* gives way to adjunct clauses with a temporal interpretation expressing simultaneity as in (12b), and, sometimes, with a modal interpretation as in (12c):

(12) a. [*gure etxe-ra etorriko d-en /*d-ela*] *lagun-a ez duzu ezagutzen*
our home-ALL come AUX-COMP AUX-COMP friend-ART NEG AUX know
'You do not know the friend that will come to our house'
b. [*gure-gana datorr-ela /*datorr-en*], *kotxe bat-ek kolpatu zuen*
we-ALL come3PSG-COMP /*come3PSG-COMP car one-ERG hit AUX
'As she was approaching us, a car hit her'
c. [*zer egin ez dakida-la /*dakida-n*] *nabil*
what do NEG know.1PSG-COMP know.1PSG-COMP walk
'I go along as if I don't know what to do'

8 The symbol "??" in (11b) reflects our judgement; de Rijk (2008) gives a single "?" to similar examples. Nevertheless, the Basque Academy (Euskaltzaindia 1999; also Etxepare 2003) regards *-ela* complements without Aux-Inversion as equivalent to those with it. The majority of modern Basque texts, along with all the speakers we have consulted, do not share this view. The discrepancy in the judgments may be more apparent than real, though: given that Aux-Inversion is perceived as an innovation that affected main clauses in the language (despite its being documented since the 16th century) and that prescription identifies *-ela* complements with subordinate clauses, it follows that they should be all right without Aux-Inversion. Admittedly, the judgments for negated *-ela* complements without inversion improve as the complement gets more deeply embedded. In sum, the standard word order for regular negative *-ela* complements in modern Basque is to have Aux-Inversion; this is not so with *-en* complements and other subordinators (with the exception of relative clauses, which never allow it), where more speaker variation is found.

As for the origin of the complementizers *-en* and *-ela* a few remarks are in order. First, de Rijk (2008: 461) has indirectly connected the complementizer *-ela* with the manner affix *-ela* present in words like *no-la* 'how, in which way', deictic manner adverbials like *hon-ela* '(in) this way', *ha-la* '(in) that way', or *beste-la* '(in) some other way', by arguing they both come from the allative adposition *-ra* 'to'; other experts in the history of Basque such as Lakarra (p.c.) agree that the connection between the manner affix *-ela* and the allative *-ra* is all the most probable (though not necessarily in the terms proposed by de Rijk), whereas the connection between the allative and the complementizer is a possible hypothesis that needs further scrutiny. With respect to the historical source of the complementizer *-en*, little is known: it has been claimed that it is related to the genitive marker *-(r)en*, but this speculation is generally rejected given that the older form of the genitive is just *-e* (Lafon 1943; Mitxelena 1977; Oyharçabal 2003).

2.2 Other complementizers with dialectal restriction

Leaving aside *-en* and *-ela*, there are four more finite verb complementizers, summarized in Table 2, with a dialectal usage: namely *-enik*, *-ena*, the dubitative or conjecture *-en*, and, finally, *bait–*, which is the only prefix form among complementizers. The chart in table 2 summarizes the dialectal distribution and the relevant grammatical information of each one of these complementizers:

Table 2

Complementizer	Dialectal distribution	Syntactic clues	Semantic clues	In non-complement clauses
-enik	Western & Central	governed by negation, matrix Q	declarative	no
-ena	Western	non-focus	factive declarative	no
-en	Navarrese, Lapurdian-Navarrese & Zuberoan	governed by certain verbs	expression of doubt, conjecture	(yes)
bait-	Lapurdian-Navarrese & Zuberoan	governed by certain predicates	factive declarative	yes

We now turn to a one-by-one review of each complementizer.

2.2.1 The Western and Central *-enik* complementizer

The complementizer *-enik*, termed "negative complementizer" by Laka (1990), is widely used in Western and Central Basque and its distribution is generally determined by a negated matrix verb, as is shown in (13a) and (13b), by an inherently negative verb as in (13c) or, much less frequently, by a matrix question as in (13d):

(13) a. *Ez dut esan* [*Jon etorri d-enik*]
NEG AUX say Jon come AUX-COMP
'I didn't say that John came'

b. *Gure alabak ez du sinesten* [*lurra eguzkia-ren inguruan dabil-enik*]
our daugther:ERG NEG AUX believe earth sun-GEN around walk-COMP
'Our daughter doesn't believe that the earth revolves around the sun'

c. *Bi udaltzain-ek ukatu zuten* [*bidaia-n Rubio-ren bizkartzain izan zir-enik*]
two police_officer-PL.ERG deny AUX trip-LOC Rubio-GEN bodyguard be AUX-COMP
'The two police officers denied that they had been Rubio's bodyguards during the trip' (*Egunkaria* newspaper 24 February 1999)

d. [*Jon etorriko d-enik*] *uste duzue?*
Jon come AUX-COMP think AUX
'Do you guys think that John will come?'

Admittedly, *-enik* clauses accept the complementizer *-ela*, as can be seen in (14a) and (14b), the *-ela* variants corresponding to examples (14a) and (14b) above:

(14) a. *Ez dut esan* [*Jon etorri d-ela*]
NEG AUX say Jon come AUX-COMP
'I did not say that John came'

b. *Gure alaba-k ez du sinesten* [*lurra eguzkia-ren inguruan dabil-ela*]
our daugther-ERG NEG AUX believe earth sun-GEN around walk-COMP
'Our daughter doesn't believe that the earth revolves around the sun'

Nevertheless, all grammarians (Laka 1990; Uribe-Etxebarria 1994; Euskaltzaindia 1999) agree that there is a basic interpretive difference: *-enik* clauses are necessarily interpreted under the scope of negation, whereas the same is not true in the case of *-ela*. Put in other words, the interpretation of *-enik* clauses usually corresponds to the Spanish polarity subjunctive triggered by negation or interrogation among others (Bosque 2012), whereas *-ela* clauses correspond to the Spanish indicative, as the following Spanish translations of (13b) and (14b) indicate.

(15) a. *Mi hija no cree [que la tierra gire alrededor del sol]*
my daughter not believes COMP ART earth revolve.SBJV around of.ART sun
'My daughter doesn't believe that the earth revolves around the sun'
b. *Mi hija no cree [que la tierra gira alrededor del sol]*
my daughter not believes COMP ART earth revolve.IND around of.ART sun
'My daughter doesn't believe that the earth revolves around the sun'

That the earth revolves around the sun is not taken for granted in the Spanish (15a) example with subjunctive mood or the Basque example (13b) with *-enik*, but the speaker assumes that the Galilean view is a true fact in the (15b) example with indicative and the Basque (14b) example with *-ela*. This may give rise to interesting contrasts, as we will show in Section 3.1.

As for the origin of this complementizer, its shape calls for an obvious explanation: *-enik* is likely to be the concatenation of the complementizer *-en* and the partitive polarity determiner *-ik*, with which it shares distribution (cf. de Rijk 1972):

(16) a. *Ez daukat diru-rik*
NEG have money-PART
'I don't have any money'
b. *Ba-daukazu diru-rik?*
BA-have money-PART
'Do you have any money?'

However, the partitive determiner is common to all dialects of Basque, and *-enik* is restricted to just two dialects; therefore a bimorphemic (syntactic) analysis of this complementizer is far from obvious from a synchronic perspective.

2.2.2 The Western complementizer *-ena*

This complementizer is restricted to Western Basque. The range of verbs that may take *-ena* declarative complements includes verbs like *aitortu* 'confess', *erakutsi* 'show', *aditu*/*entzun* 'hear', *igarri* 'notice', *jakin* 'know', *argi egon* 'be clear', *gogoratu* 'remember', *konturatu* 'realize', *seguru izan/egon* 'be sure', *ezagun izan* 'be obvious' and so on, many of which are factive verbs. Grammarians (Azkue 1923; Arejita 1984; Euskaltzaindia 1999; Artiagoitia 2003) agree that, whenever *-ena* is used, the speaker generally takes for granted the truth value of the proposition expressed by the complement clause; in other words, *-ena* is a kind of factive marker. We give a couple of examples in (17), adding the corresponding examples with *-ela*, which is also possible:

(17) a. {*ba-dakit, igarri, entzun}* *dut* [*Jon etorri d-ena*]
ba-know guess hear AUX Jon come AUX-COMP
'I {already know, guessed, heard} that John came'
b. *Ezagun da* [*gaztea zar-ena*]
obvious is young are-COMP
'It's obvious that you are young'
(Euskaltzaindia 1999: 68)
c. *Argi dago* [*Amaia-k asko ikasten du-ena*]
clear is Amaia-ERG much study AUX-COMP
'It's clear that Amaia studies a lot'
(Artiagoitia 2003: 646)

(18) a. {*ba-dakit, igarri, entzun*} *dut* [*Jon etorri d-ela*]
ba-know guess hear AUX Jon come COMP-AUX
'I {already know, guessed, heard} that John came'
b. *Ezagun da* [*gaztea zar-ela*]
obvious is young are-COMP
'It's obvious that you are young'
(Euskaltzaindia 1999: 68)
c. *Argi dago* [*Amaia-k asko ikasten du-ela*]
clear is Amaia -ERG much study AUX-COMP
'It's clear that Amaia studies a lot'

For those speakers who use both *-ena* and *-ela* as complement clauses to this type of verbs, there seems to be a slight difference in the factual interpretation of the complement clause: the choice for *-ena* leads to presuppose the truth of the complement clause, whereas *-ela* sentences are more neutral in this respect; that is, they do not prompt any nuance on the truth value of the complement proposition.

Given that informational focus provides a new, non-presupposed, piece of information, the complementizer *-ena*, due to the fact that it always presupposes the truth of the proposition, is incompatible with an information focus interpretation of the complement clause, as is shown in (19):

(19) Q: *Zer igarri duzu?*
what guess AUX
'What did you guess?'
A: *Jon etorri d-ela* */*Jon etorri d-ena*
Jon come AUX-COMP Jon come AUX-COMP
'That John came'

And, as such, it does not occupy the preverbal focus position (cf. Euskaltzaindia 1999; Artiagoitia 2003).

The morphological shape of *-ena* is also fairly transparent: it seems to be composed of the complementizer *-en* and the singular article *-a*; but, once again,

this does not imply it should be treated as bimorphemic from the point of view of present Basque.[9]

2.2.3 The non-interrogative *-en* complementizer in eastern Basque

As for the third complementizer, Lapurdian-Navarrese, Zuberoan and Navarrese *-en* is homophonous with the general *-en* present in wh-interrogatives and exclamatives, but its syntactic distribution is distinct and its use restricted to eastern dialects of Basque. As far as we know, its syntactic distribution has not been well studied: it is lexically restricted to complement clauses of certain predicates such as *uste izan* 'think', *baditake* 'it is possible', and *iduri* 'seem' (Lafitte 1944; Euskaltzaindia 1999; Camino 2004; Epelde 2003), whether the matrix verb is negated or not. Here are some relevant examples, taken from the Lapurdian-Navarrese, (20a), (20b), (20c) and (20d) and the Zuberoan, (20e), dialects:

(20) a. *Ba-dita-ke* [*ez-t-en* *hain* *gaztia*]
ba-AUX-can NEG-is-COMP so young
'It could be that he is not so young'
(Epelde 2003: 194)

b. *iduri* *zaut* [*"marluza"* *err(a)ten* *d-en*]
seem AUX "marluza" say AUX-COMP
'It seems to be that it is said "marluza"'
(Camino 2004: 477)

c. *etzaut* *iduitzen* [*bera* *lotzen* *ahal* *d-en*]
NEG.AUX seem it burn can AUX-COMP
'It does not seem to me that it can burn by itself'
(Camino 2004: 477)

d. *Nik* *uste* [*hortarik* *hasi* *dir-en*...] *zerbait-etarik* *hasteko*
I-ERG think there-ABL start AUX-COMP something-ABL start-to
'I think they started from there just to start from somewhere'

e. *üdüritzen* *zitazüt* [*gaizak* *gaizki* *eginik* *izan* *dir-en*]
seem AUX things bad do be AUX-COMP
'It seems to me that things have been done badly'
(Bedaxagar, p.c.)

A more appropriate English translation of (20b) should be 'It seems to me that it might be said "marluza" (the word for 'hake')'. Needless to say, all dialects may use *-ela* in the contexts in (20), and, with the corresponding nuances, Central and Western Basque might use *-enik* in the examples where the matrix verb is negated (= 20c).

9 For a treatment of *-ena* as a truly bimorphemic complementizer in synchronic terms along the lines of Kiparsky and Kiparsky (1970), see, for instance, Eguzkitza (1997).

2.2.4 The complementizer *bait-* in the northern dialects

Finally, we will touch on the complementizer *bait-*, the only one that is a prefix, rather than a suffix, on finite verb forms. Although the basic form is *bait-*, its morphological shape can change depending on the verb-initial consonant and be pronounced and written as simply *bai-* (with devoicing of a following stop) or even *peit* after *ez* 'not' as in (21a). This complementizer has a very limited distribution in northern Basque (in the Lapurdian-Navarrese and Zuberoan dialects); it heads complements to happening verbs like *gertatu* 'happen', *gutitarik egin* 'do by little' and to emotive predicates like *domaia izan* 'be a pity' as in (21a), and *harrigarri, bitxi izan* 'be surprising' as in (21b); it can also introduce a declarative complement referring to a presentational demonstrative as in (21c):

(21) a. *Domaia da* [*ez-peit-a jinen*]
pity is NEG-COMP-AUX come
'It's a pity that he will not come'
(Epelde 2004: 195)

b. *Bitxi da* [*ez bait-uzu sekula ho-lako-rik ikusi*]
strange is NEG COMP-AUX never that-like-DET see
'It's strange that you have never seen anything like that'
(Epelde 2004: 195)

c. *Hau da haren abantailik handien-a* [*ez bait-u ainitz xahutzen*]
this is its advantage biggest-ART NEG COMP-AUX much spend
'This is its main advantage, that it does not spend much'
(Lafitte 1944: 403)

Although this requires further research, the *bait-* complements also have a factive flavour, in the sense that they express propositions presupposed as facts.[10]

We will now present a feature value analysis of *-en* and *-ela*.

---

10 An anonymous reviewer wonders about the situation in other dialects: these usually resort to non-finite complements (in the form of nominalizations of the type seen in (6a) or to *-ela* complements. In fact, speakers of the northern dialects can also alternate *bait-* complement sentences with nominalizations such as (i) below:

(i) {*Bitxia, pena, harrigarria*} *da Jon ez etor-tzea*
strange pity surprising is Jon NEG come-NMLZ
'It is {strange, a pity, surprising} for John not to come'

(ii) *Hau da haren abantailarik handiena, ez du-ela asko gastatzen.*
This is its advantage biggest:ART NEG AUX-COMP much spend
'This is its main advantage, that it does not spend much'

3 Towards a feature analysis of *-en* and *-ela*

As a first rough approximation, it would seem that one can define *-en* as a wh-complementizer, one which is the spell-out of the feature [wh] in Comp selected by a matrix verb with a +[wh] feature in its subcategorization frame, whereas *-ela* should be defined as simply declarative. This kind of characterization goes back to Goenaga's (1985) work, sketched in (22):

(22) *-en* = +wh, +finite
-ela = −wh, +finite
(Goenaga 1985: 506)

However, three reasons move us to try a finer characterization of these two basic complementizers; first, the fact that *-en* is present in other kind of subordinate clauses as seen before (relatives and also comparatives); second, the fact that even for a simple characterization of the complementizer system, more than one single feature is often used. For example, Rizzi (1990: 67) used the features [±wh, ±predicative] for a characterization of (English) complementizers. Interestingly, relative clauses without relative pronouns were characterized as [−wh, +predicative]; these are precisely the most typical relative clauses in Basque, and require the presence of *-en*, as seen in examples (5c) and (12a) above, repeated here as (23a) and (23b):

(23) a. *Jon-ek*$_i$ [e_i *erosi* *du-**en***] *kotxe-a* *erakutsi* *digu* (= 5c)
Jon-ERG buy AUX-COMP car-ART show AUX
'John has shown us the car that he has bought'
b. [*gure* *etxe-ra* *etorriko* *d-en*] *lagun-a* *ez* *duzu* *ezagutzen* (= 12a)
our home-ALL come AUX-COMP friend-ART NEG AUX know
'You do not know the friend that will come to our house'

Thus, a simple [+wh] might not be adequate for a characterization of the complementizer *-en*. The third reason is that, in current theoretical syntax, privative features are generally preferred over binary valued features, since the latter may often allow too many possibilities (Adger 2003: 30); so wherever possible, we would like to use the [wh] feature as a privative feature, and not as a binary valued feature. The task is complicated given that interrogative complements are generally taken to involve a Q-feature *and* a wh-feature (cf. Rizzi 1996, Adger 2003).

What we seem to need, then, is a feature that will reflect that *-en* generally spells out a complementizer that agrees with an operator in its specifier; this operator may be a wh-phrase in the case of indirect questions and wh-exclamatives, or a null (wh) operator in the case of yes/no indirect questions, relatives,

and comparatives.[11] (24) is a summary of operator-constructions, all of them occurring with *-en*:

(24) a. wh-question: [*nor* etorri d-*en*] galdetu dut
who come AUX-COMP ask AUX
'I asked who came'

b. wh-exclamative: harritzekoa da [*nor* etorr d-*en*]
surprising is who come AUX-COMP
'It's surprising who came'

c. yes/no indirect question: [*Op* Jon etorri d-*en*] galdetu dut
Jon come AUX-COMP ask AUX
'I asked whether John came'

d. relatives: [*Op* etorri d-*en*] gizon-a...
come AUX-COMP man-DET
'The man that came...'

e. comparatives: [*Op* esan duzu-*n*] baino jende gehiago
say AUX-COMP than people more
etorri da
come AUX
'More people came than you said'

One tentative solution would be to use the feature [operator] without the exclusion of [±wh], since not all operators have the feature [+wh]. In this case, we would replace (22) with (25):

(25) *-en* = Comp. [operator, ±wh], [finite]
-ela = Comp. [declarative], [finite] (or simply = default Comp. [finite])

In other words, *-en* would be the spell-out of a complementizer agreeing with an operator, be it marked [wh] or not, sitting in its specifier.

Alternatively, if we adopt for Basque the standard assumption that the yes/no operator of direct and indirect questions is endowed with a [wh] feature that may trigger I-to-Comp movement (e.g., Adger 2003), a slight modification of Goenaga's original characterization would suffice for the purposes of accounting for the presence of *-en* in complement clauses. The relevant modification is in (26):

(26) *-en* = [wh], [finite]
-ela = [declarative], [finite] (or simply = default Comp. [finite])

11 That Basque relative clauses involve null operators has been argued for by Oyharçabal (1988) and Artiagoitia (1992); Ortiz de Urbina (1995, 1999) has extensively argued for the existence of yes/no operators in questions; finally, the analysis of null operators in comparatives is due to Sáez (1989).

That is, *-en* would be the spell-out of a wh-complementizer selected by the matrix verb and agreeing with a wh-operator (overt or covert) in its specifier.[12]

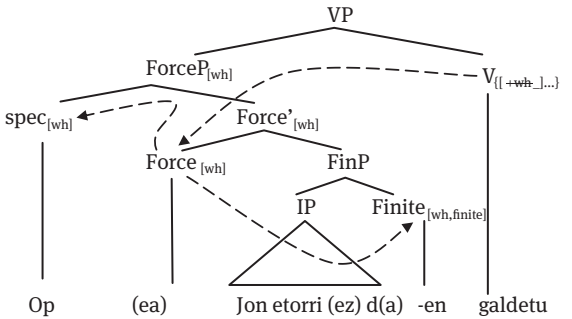
A more fine-grained analysis of the complementizer system in Basque is proposed in Ortiz de Urbina (1999). Adopting Rizzi's (1997) split CP system, he proposes to analyze the structure of a Complementizer Phrase as in (27), where *-en* and *-ela* would be associated with finiteness, and not with force, given that they are affixal subordinators present only in finite clauses:

(27) $[_{\text{ForceP}}$ *Force* $[_{\text{TopP}}$ Topic $[_{\text{FocP}}$ Focus $[_{\text{FiniteP}}[_{\text{IP}}$ NP VP INF] Finite]]]]

With respect to interrogative complements, exemplified in (28), Ortiz de Urbina (1999) proposes that the complementizer *-en* spells out the [wh] feature in Force, which is itself selected by the matrix V.[13] According to Ortiz de Urbina, the yes/no operator would be in the specifier of Force, in which case the feature agreement between the operator and the complementizer *-en* would only be indirect, mediated by the *Force-Finite* head selectional relationship.[14] The tree diagram in (29) illustrates how the analysis works:

(28) [*(ea)* *Jon* *etorri* *(ez)* *d-en*] *galdetu* *dute*
ea Jon come NEG AUX-COMP ask AUX
'They asked whether John {came, did not come}'

(29)



In the following section, we will tackle what we feel are two interesting problems that bear on the characterization presented in (26) and (27) above: the use of *-enik*

12 Then, if the null operators present in non-complement clauses such as relatives and comparatives can also be said to be [wh], we would reach a near unified account of *-en* as a subordinator.
13 Leftward movement of wh-phrases and the Inflection-Finite complex to Focus-Phrase would depend on the presence of a non-interpretable wh-feature (strong in Basque) in an overt wh-phrase or similar operators (e.g., negation).
14 One reviewer points out that "the idea of identifying (illocutionary) forces in dependent clauses runs counter to a widespread idea (advocated by Searle and, more recently Cristofaro

in Western and Central Basque, and finally, the co-occurrence of both *-en* and *-ela* in the so called subjunctive complements.

4 Open questions in the Basque complementizer system

4.1 Polarity, factivity and finite complementizers

We saw above that Western and Central Basque have a third Comp. *-enik*, argued to be a "negative complementizer" by Laka (1990) because its distribution is generally determined by a negated matrix verb, e.g. in (30a) and (30b), although it can also show up under the scope of a matrix question (e.g., in 30c); (30d) shows that *-enik* is only licensed in exactly those two environments:

(30) a. *Ez dut {esan /uste} [Jon etorri d-enik]*
NEG AUX say /think Jon come AUX-COMP
'I {didn't say/ don't think} that John has come'
(cf. Spanish *No he dicho que Jon haya venido*)

b. *Ez dut jakin [Jon etorri d-enik]*
NEG AUX know Jon come AUX-COMP
'I did not know that John has come'
(cf. Spanish *No sabía que Jon hubiera venido*)

c. *Entzun duzu [Jon etorri d-enik]?*
hear AUX Jon come AUX-COMP
'Did you hear that John came?'
(cf. Spanish *Has oído que Jon haya venido?*)

d. **Entzun dut [Jon etorri d-enik]*
hear AUX Jon come AUX-COMP
'I heard that John came'
(cf. Spanish **He oído que Juan haya venido*)

Example (31) illustrates that *-enik* is not possible in embedded negative or question clauses if the matrix sentence is declarative:

2003) according to which subordinate clauses are defined exactly as being without illocutionary force." His suggestion is to "maintain that subordinate clauses may have illocutionary VALUES, but do not have illocutionary FORCES." Even if one agrees with the suggestion, it must be noted, however, that Rizzi's (1997) position is to propose a syntactic head *Force* to encode morphosyntactic features related to illocutionary force; thus, the lack of illocutionary force would also be coded on this syntactic head. But see Heycock (2006), Haegeman (2006a, 2006b), Wiklund et al. (2009) for the view that certain complement clauses behave like root clauses and may have illocutionary force.

(31) a. **Uste dut [Jon ez d-enik etorri]*
think AUX Jon NEG AUX-COMP come
'I think that John has not come'
b. **[Jon etorri d-enik] jakin nahi zuten*
Jon come AUX-COMP know want AUX
'They wanted to know that John has come'

Laka claims that only negative complementizers can locally license Negative Polarity Items (NPIs) in complement clauses; NPIs are not licensed in complement clauses headed by *-ela*. The contrast is shown in (32), where *inor* 'anybody' is the NPI to be licensed:

(32) a. *Ez dut esan [inor etorri d-enik]*
NEG AUX say anybody come AUX-COMP
'I didn't say that anyboy came'
b. **Ez dut esan [inor etorri d-ela]*
NEG AUX say anybody come AUX-COMP
'I didn't say that anybody came'

Uribe-Etxebarria (1994), on the other hand, argues that the licensing of NPIs is independent of the existence of a negative $\text{Comp}_{[+\text{neg}]}$. She provides data where NPIs are licensed despite the absence of a negative complementizer, and derives the contrast in (32) above from the fact that NPI licensing takes place at LF. On this view, the embedded *-ela* clause raises above the matrix negation at LF, leaving the NPI unlicensed, whereas the *-enik* subordinate clause must remain within the scope of the matrix clause, [15] just like the Romance polarity subjunctive.[16]

---

15 In the dialects which do not have *-enik* the contrast does not hold and NPIs are possible in complement clauses without *-enik*, provided they are under the scope of matrix negation or question:

(i) *Ez dut uste nehor enoatu d-**en**.*
NEG AUX think anybody get_bored AUX-COMP
'I do not think anybody got bored'

(ii) *Ez daukagu ere gure erran-a nehor-k gaizki hartu behar luke-**la**.*
NEG have even our saying-ART anybody-ERG badly take have AUX-COMP
'We do not think that anybody should take our saying in a bad way'

For these dialects, it seems as though raising of the complement is not needed.

16 Polarity subjunctive is so called because its licensing resembles that of Negative Polarity Items in that it is also triggered by negation (cf. [32a] above) or questions (cf. example [ia] below); the same subjunctive is banned in regular affirmative sentences (cf. [ib]):

(i) a. *¿Recuerdas que dijera algo con sentido?*
remenber that said anything with sense
'Do you remember that he said [SBJV] anything meaningful?'
(Bosque 2012: 375)

In relation to this, Laka (1990) and Uribe-Etxebarria (1994) both agree that selecting one complementizer over the other leads to a different semantic interpretation associated with presupposition. The contrast is provided in (33):

(33) a. *Ez dut esan [ama etorri d-enik]*
NEG AUX say mother come AUX-COMP
(= no presupposition implied)
'I didn't say that mom came'
b. *Ez dut esan [ama etorri d-ela]*
NEG AUX say mother come AUX-COM
(= truth of subordinate taken for granted)
'I didn't say that mom came (and in fact she did)'

It is well known since Kiparsky and Kiparsky (1970) that the semantic differences between factives and non-factives might be reflected in the syntax: they proposed that factives are syntactically more complex and involve a silent "fact" noun with a sentential complement.

However, in the recent literature there have been proposals which hold the opposite view (see Nichols 2001; McCloskey 2005; Haegeman 2006; de Cuba 2007). In short, these proposals all claim that the presupposed status of factive complement clauses is the default interpretation and that, instead, what requires explanation is why certain CPs have a non-presupposed (non-factive) interpretation. According to this view, non-factive complements are structurally more complex: they project an extra functional layer which is absent in factive complements. In particular, de Cuba (2007) proposes that this extra projection has a semantic operator which 'removes the speaker from responsibility for the truth content of the embedded clause' (2007: 9). This is precisely the interpretation that Kempchinsky (2009) attributes to the so-called polarity subjunctive (which she formalizes by endowing Force with a W_{su} feature), and this is also the interpretation we get with the complementizer *-enik*.

Thus, assuming the structure given in (28), partially repeated in (34a), we suggest analyzing Basque *-enik* as the realization of Finite° when a non-factive complement clause is triggered by a predicate in connection with a matrix polarity operator. The null operator of the complement clause would provide the non-presupposed interpretation characteristic of these clauses. The structure of a clause with *-enik* is illustrated in (34b):

b. **Recuerdo que dijera algo con sentido.*
remember that said anything with sense
'I remember that he said [SBJV] something meaningful'

(34) a. $[_{\text{ForceP}}$ *Force*$_{[\text{Wsu}]}$ $[_{\text{FiniteP}}$ $[_{\text{IP}}$ NP VP INF] *Finite*]]
b. neg V $[_{\text{ForceP}}$ Op *Force*$_{[\text{Wsu}]}$ $[_{\text{FiniteP}}$ $[_{\text{IP}}$ NP VP INF] *-enik*$_{[\text{Wsu}]}$]]

Contrary to what happens in Romance (where the verbal Mood inside IP identifies the evaluation feature), it would be the head Finite outside IP the one that identifies the W feature in Force. The feature analysis of *-enik* would then be something like (35):

(35) *-enik* = [W_{su}], [finite]

Where W_{su}, with the subindex$_{su}$, is the feature proposed by Kempchinsky (2009), and roughly signals that the propositional content of the complement is to be evaluated in the context of the matrix subject's beliefs.

4.2 The choice of *-en* vs. *-ela* in subjunctive complement clauses

Subjunctive has been defined as the realization of a specific mood usually associated with intentionality, volition, command or world evaluation. With respect to Basque, it also exhibits a particular syntactic behavior: first, it always occurs in dependent clauses introduced by an overt complementizer; and second, it occurs with a different set of auxiliaries. These auxiliaries are **edin* for intransitive predicates and **ezan* for transitive predicates, both of which are reconstructed verb infinitives. The following examples in (36) illustrate how the auxiliary shifts from *izan* 'be' in ordinary declarative sentences (e.g., in [36a]) to **edin* in subjunctive complement clauses as in (36b); and how the auxiliary **edin* cannot be a bare form in (36c) unless the modal suffix *-ke* is added in (36d):

(36) a. *Jon* *etorri* *da*
Jon come be.AUX
'John came'
b. [*Jon* *etor* *dadi-n*] *nahi* *dut*
Jon come edin.AUX-COMP want AUX
'I want that John come'
c. **Jon* *etor* *dadi*
Jon come edin.AUX
'John comes'
d. *Jon* *etor* *dai-teke*
Jon come edin.AUX-FUT
'John may come'

As for the selection of subjunctive complements, we adhere to the traditional idea that subjunctive in Basque is lexically selected by desiderative/volitional verbs,

directive or effective (influence/command) verbs, reported speech verbs, and opinion-predicates (Ormaechevarria 1959; Goenaga 1980, 1997; Euskaltzaindia 1999). Examples in (37) summarize the usual range of possibilities:

(37) a. [*Jon etor dadi-n*] {*nahi, behar, espero,...*} *dut*
Jon come edin.AUX-COMP want need hope AUX
'I {want, need, hope} that John come'

b. [*Jon etor dadi-n*] {*erregutu, eskatu, debekatu, agindu*} *dut*
Jon come edin.AUX-COMP beg demand forbid order AUX
'I {begged, demanded, forbade, ordered} that John come'

c. [*Jon etor dadi-n*] {*adierazi, esan*} *dizut*
Jon come edin.AUX-COMP express say AUX
'I {expressed, said} that John come'

d. {(*Ona, zilegi, beharrezkoa*) *da /ondo iruditzen zait*} [*Jon etor dadi-n*]
good licit necessary is well seem AUX Jon come edin.AUX-COMP
'{It's {good, licit, necessary} / it seems ok to me} that John come'

Contrary to Romance, we don't find subjunctives as complements to emotive-factive predicates or predicates expressing doubt, as can be seen in (38):

(38) a. *[*Jon etor dadi-n*] *deitoratzen dut*
Jon come edin.AUX-COMP regret AUX
'I regret that John come'

b. *[*Jon etor dadi-n*] *duda egiten dut*
Jon come edin.AUX-COMP doubt do AUX
'I doubt that John come'

From the point of view of this article, the crucial point is that many subjunctive complements may alternate between complementizers *-en* and *-ela*, with no apparent semantic difference. To be more precise: *-en* is the universal or default complementizer in subjunctive clauses, but as a matter of dialect (and diachronic) variation, *-ela* is becoming more and more common in southern dialects as complement to volitional verbs as in (39a) and, above all, directive verbs (39b); the exclusive use *-en* is restricted to opinion predicates as in (39d). Relevant data are in (39):

(39) a. [*Jon etor dadi-la*] {*nahi, behar, espero,...*} *dut* (cf. 37a)
Jon come edin.AUX-COMP want need hope AUX
'I {want, need, hope} that John come'

b. [*Jon etor dadi-la*] {*erregutu, eskatu, debekatu, agindu*} *dut* (cf. 37b)
Jon come edin.AUX-COMP beg demand forbid order AUX
'I {begged, demanded, forbid, ordered} that John come'

c. [*Jon etor dadi-la*] {*esan, adierazi*} *dizut* (cf. 37c)
Jon come edin.AUX-COMP {say, espress} AUX
'I {expressed, said} that John come'

d. **{(Ona, zilegi, beharrezkoa) da / ondo iruditzen zait}* [*Jon etor dadi-la*] (cf. 37d)
good licit necessary is well seem AUX Jon come edin.AUX-COMP
'{It's {good, licit, necessary}/it seems ok to me} that John come'

The question is: if our proposal is that *-en* spells out a [wh]/operator and [finite] complementizer and *-ela* is a declarative or default [finite] complementizer, how do we account for the occurrence of both complementizers in subjunctive clauses?

As for the default use of *-en* as a subjunctive complementizer, it does have a fairly reasonable explanation in the account of subjunctive complements developed by Kempchinsky (2009). Taking as departure the observation that lexically subjunctive complements display the Disjoint Reference Effect whereby the matrix subject and the embedded subject of the subjunctive clause cannot co-refer,[17] this author regards subjunctive complements as hidden imperatives and proposes that the obviation effect is due to a quasi-imperative operator located in the head of Finite-Phrase which induces the interpretation 'anyone other than the matrix subject'. Kempchinsky (2009: 1800) proposes the structure in (40) for a selected subjunctive complement:

(40) ...V_w [$_{CP}$ [$_{ForceP}$ Force$_{[uW]}$] [$_{FinP}$ [Fin$_{[uW]}$ Op] [$_{IP}$ (DP) [$_{MoodP}$ [V+TM $_W$] [$_{TP}$ ...]]]]

Selection checking (Agree)

The matrix verb selects for a clause type bearing the (uninterpretable) feature World $_{[uW]}$, which roughly means that the propositional content expressed by the subjunctive is being evaluated by the matrix subject, and hence represents selection of subjunctive mood. In a language with mood distinction, the verbal complex is the one endowed with the corresponding interpretable feature that can check and delete the feature present in Force and also in Finite.

If we extend this analysis to Basque subjunctives, it turns out that the complementizer *-en* is expected, given that on this approach the head Finite hosts

17 For example:

(i) *pro*$_i$ *quiere que pro*$_{\{*i/j\}}$ *venga*
he$_i$ wants that he$_{\{*i/j\}}$ come[SBJV]
'he$_i$ wants for him$_{\{*i/j\}}$ to come'

The matrix subject and the embedded one are necessarily different people.

a quasi-operator element. Thus, either the previous feature analysis in (26) (repeated here) will do:

(26) *-en* = Comp. [operator, ±wh], [finite]
-ela = Comp. [declarative], [finite] (or simply = default Comp., [finite])

Or else, we can change (26) to the slightly more sophisticated version in (41):

(41) *-en* = {[wh]/[operator]}, [finite]
-ela = [declarative], [finite] (or simply = Comp. [finite])

In other words, the common ground of interrogatives and subjunctive complements in Basque is the presence of an operator feature in Comp or in Finite, be it a wh-operator or not.[18]

Now, one non-trivial question is pending: what makes it possible to use *-ela* in many subjunctive clauses, clearly non-declarative, on a par with *-en*? [19] We honestly do not have a sharp answer to this question, nonetheless we wish to make a couple of remarks that may give us a clue for future research:

a. As pointed out by Goenaga (1997), *-ela* has spread over the years to subjunctive complements in the case of *communication* verbs, i.e. indirect speech verbs and directive verbs; much less in the case of volitional verbs, and not at all in the case of opinion predicates. Consequently, one interesting approach would be to identify *-ela* with some acquired feature, say informally [quotative], and propose that subjunctive complements may alternate between an operator-like finite head and a communication-oriented finite head, clearly absent in opinion predicates. The drawback of this approach is that we are led to abandon Kempchinsky's (2009) main insight, namely, that subject obviation in subjunctives is across the board due to the presence of an operator. So this approach would just be a mere reformulation of the data in different terms.
b. We think that a more interesting suggestion is, however, possible: namely, that subjunctives with *-ela* do in fact also contain the complementizer *-en*; in other words, we suggest that subjunctives may surface just with the operator complementizer *-en* or, alternatively, with both *-en* and *-ela*, collapsed into a

18 It is interesting to note that the obviation effect of subjunctives does not hold in non-selected polarity subjunctives (cf. *no creo que llegue a tiempo* 'I don't think I will arrive on time'), where the operator in Finite is absent; but similar Basque examples never use subjunctive complements.

19 A reviewer suggests that analyzing Basque subjunctives as SoA markers predicts that *-ela*, an element which takes scope over propositions, will be compatible with subjunctives, given that propositions necessarily involve SoAs (but not vice versa). The problem is that not all subjunctive complements allow *-ela* (cf. [39d] and Goenaga 1997 on this issue).

single *-ela*. Let's see what this implies. Due to the morphophonemics of *-ela* (only *-la* when attached to consonant-ending verbal forms), every time it is attached to a past form ending in the suffix *-n*, the nasal consonant disappears. We show this in (42):

(42) a. *etorri* *nintze-n*
come 1PSG-PST
'I came'
b. *etorri* *nintze-n-la* > *etorri nintzela , *etorri nintzenla*
come 1PSG-PST-COMP
'that I came'

Thus, it just happens that if we attach the complementizer *-ela* to a subjunctive form that already has the complementizer *-en*, the result will be for the final /n/ consonant to drop, as in (43):

(43) *etor* *n-adi-n* + *la* > *etor nadila, *etor nadinla*
come 1PSG-edin.AUX-COMP COMP
'that I come'

It is therefore perfectly plausible that we have both *-en* and *-ela* in the examples in (39) above, with *-en* not being realized for morphophonemic reasons.

If the suggestion is on the right track, then one could maintain that *-en* in subjunctive complements represents the head Finite which contains, just like in interrogative complements, an operator element; subjunctive *-ela* could then be perhaps identified with the (higher) head Force, representing clause type, possibly associated with the informal feature *quotative*. Given that opinion predicates are not communication verbs, it is to be expected that they will never take *-ela*. Admittedly, this amounts to the proposal that there are in fact two distinct *-ela* complementizers: one that is the default finite or Comp. head in declaratives (see [24] or [39] above); and a *quotative* one with true (communication) content in subjunctive complements, used in conjunction with *-en*. The latter *-ela*, generated in a head initial Force, would possibly be cliticized onto the verb in the phonological component.

5 Conclusions

In this article we hold that Basque supports the view that complementizers have distinct semantic values related to truth value, factivity, mood and evaluation. We mainly focus on the two most general and productive finite complementizers (*–ela* and *-en*), but the discussion extends to other complementizers which show a

more restricted dialectal distribution, given that precisely those complementizers shed more light on the semantic values associated with them. More specifically, we claim that:

a. The basic distinction between the finite complementizers *-en* and *-ela* is a clear cut case of declarative vs. interrogative complements. However, noting that the use of *-en* extends to other types of clauses which generally involve operators (be them overt or null) such as subjunctive complement clauses, relatives, and comparatives, we have formulated the distinction in terms of the following feature analysis (see Section 3):

(44) *-en* = {[wh]/[operator]}, [finite]
-ela = [declarative], [finite] (or simply = Comp. [finite])

b. In relation to semantic values such as truth value, evaluation and factivity, we argue that there are specific complementizers (subject to dialectal distribution) which do trigger a particular interpretation on the sentential complement they head. Thus, it is shown that the Western and Central Basque *-enik* complementizer is a counterfactual marker, in the sense that it induces lack of factivity on the content of the complement clause (Section 2.2.1). Besides, it is argued that its distribution is subject to polarity contexts such as matrix negation and matrix questions (section 4.1). It is in this sense that we compare Western and Central Basque *-enik* with the Romance polarity subjunctive, as is discussed in Kempchinsky (2009) and Bosque (2012). As a matter of fact, eastern dialects which lack this complementizer make use of another complementizer, *-en*, in complements to a few verbs which express conjecture (or lack of certainty), which again suggests that complementizers are associated with a specific semantic value (Section 2.2.3). On the other side, it is argued that the Western Basque *-ena* complementizer triggers a factive interpretation on the complement sentence it heads, that is, it signals that the propositional content is judged as certain by the speaker (Section 2.2.2).

Acknowledgment

This article represents joint research and the authors are listed in strict alphabetical order. We are grateful to Petar Kehayov and Kasper Boye for having given us the opportunity to retake the issue of Basque complementation and its morphosyntax, a research topic that we have possibly neglected in Basque linguistics over the last years. Special thanks to M. Bedaxagar, I. Epelde, L. Mendiburu and J.B. Dirassar for their help with Lapurdian-Navarrese and Zuberoan data. This study

has been made possible by the research project FFI2013-46907-P, and within the research group GIC07/144-IT-210-07 and IT769-13 from the Basque Government. All errors are ours. As required by the editors, we use the term *complementizer* to refer to subordinators that introduce complement clauses. As is standard in the generative tradition, we adhere to the view that most subordinators are in fact of the same category *Complementizer*, but this view is irrelevant for the purposes of this article.

References

Adger, David. 2003. *Core syntax*. Oxford: Oxford University Press.

Aldai, Gontzal. 2008. From Ergative Case Marking to Semantic Case Marking. In Mark Donohue & Søren Wichmann (eds.), *The Typology of Semantic Alignment*, 197–218. Oxford: Oxford University Press.

Aldai, Gontzal. 2009. Is Basque morphologically ergative? Western Basque vs. Eastern Basque. *Studies in Language* 33(4). 783–831.

Arejita, Adolfo. 1984. *Euskal joskera* [Basque syntax]. Bilbao: Labayru.

Arteatx, Iñigo. 2011. *Euskarazko egitura jokatugabeak: -t(z)en morfemadun perpaus osagarrien azterketa sintaktikoa* [Basque tenseless structures: syntactic analysis of the complement clauses headed by the morpheme *-tzen*]. Leioa: University of the Basque Country PhD dissertation.

Artiagoitia, Xabier. 1992. Why Basque doesn't relativize everything. In Joseba A. Lakarra & Jon Ortiz de Urbina (eds.), *Syntactic Theory and Basque Syntax*, 11–35. San Sebastian: Supplements of ASJU.

Artiagoitia, Xabier. 2003. Complementation. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A Grammar of Basque*, 634–710. Berlin: Mouton de Gruyter.

Azkue, Resurrección María de. 1969 [1923]. *Morfología Vasca*, 2nd edition. Bilbao: La Gran Enciclopedia Vasca.

Baxok, Erramun. 2007. *Euskararen egoera Ipar Euskal Herrian* [The situation of the Basque language in the North Basque Country]. Baiona: Euskararen Erakunde Publikoa.

Bosque, Ignacio. 2012. Mood: Indicative vs. Subjunctive. In José Ignacio Hualde, Antxon Olarrea & Erin O'Rourke (eds.), *The Handbook of Hispanic Linguistics*, 373–394. Chichester: Wiley-Blackwell.

Camino, Iñaki. 2004. Nafarroa Behereko euskara [The Basque of Low Navarre], *Fontes Linguae Vasconum* 97. 445–486.

Chomsky, Noam. 1977. On Wh-movement. In Peter Culicover, Thomas Wasow and Adrian Akmajian (eds.), *Formal Syntax*, 71–132. New York: Academic Press.

Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.

Croft, William. 1990. *Typology and Universals*. Cambridge: Cambridge University Press.

De Cuba, Carlos. 2007. *On (non)factivity, clausal complementation and the CP-field*. Stony Brook University PhD dissertation.

Dryer, Matthew S. 2007. Word order. In Timothy Shopen (ed.), *Clause Structure, Language Typology and Syntactic Description,* vol.1, 2nd Edition, 61–131. Cambridge: Cambridge University Press.

Duguine, Maia. 2012. Basque nominalizations and the role of structural Case in the licensing of null arguments. In Urtzi Etxeberria, Ricardo Etxepare and Mryriam Uribe-Etxebarria (eds.), *Noun Phrases and Nominalization in Basque,* 333–374. Amsterdam: John Benjamins.

Eguzkitza, Andolin. 1997. Indirect questions and other adnominals in the grammar of Basque. In B. Palek (ed.), *Proceedings of LP'96*, 332–350. Prague: Charles University Press.

Elordieta, Arantzazu. 2001. *Verb Movement and Constituent Permutation in Basque* (LOT series 47). Utrecht: Utrecht University.

Elordieta, Arantzazu. 2002. The Role of Verbal Agreement in Licensing Null Arguments. In Pablo Albizu, & Beatriz Fernandez (eds.), *Kasu eta Komunztaduraren gainean. On Case and Agreement*, 113–132. Leioa: University of the Basque Country.

Epelde, Irantzu. 2003. *Larresoroko euskara* [The Basque variety of Larresoro]. Leioa: University of the Basque Country PhD dissertation.

Etxepare, Ricardo. 2003. Negation. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A Grammar of Basque*, 516–564. Berlin: Mouton de Gruyter.

Euskaltzaindia. 1985. *Euskal Gramatika. Lehen Urratsak–I* [Basque grammar. First steps-I]. Bilbao: Euskaltzaindia.

Euskaltzaindia. 1999. *Euskal Gramatika. Lehen Urratsak–V* [Basque grammar. First steps-V]. Bilbao: Euskaltzaindia.

Euskaltzaindia. 2005. *Euskal Gramatika. Lehen Urratsak–VI* [Basque grammar. First steps–VI]. Bilbao: Euskaltzaindia.

Eusko Jaurlaritza. 2009. *IV Mapa soziolinguistikoa.*[The 4th sociolinguistic map]. Vitoria: Eusko Jaurlaritza.

Goenaga, Patxi. 1980. *Gramatika Bideetan* [Along the paths of grammar]. 2nd edition. Donostia: Erein.

Goenaga, Patxi. 1984. *Euskal sintaxia: konplementazioa eta nominalizazioa* [Basque syntax: complementation and nominalization]. Leioa: University of the Basque Country PhD dissertation.

Goenaga, Patxi. 1985. Complementación y nominalización en euskara. *International Journal of Basque Linguistics and Philology* 19(2). 495–570.

Goenaga, Patxi. 1997. Subjunktiboaz eta aginteraz zenbait ohar: *-en* eta *-ela* atzizkien erabilera [Some notes on the subjunctive and the imperative: the use of *-en* and *-ela* suffixes]. In Ricardo Badiola (ed.), *Federiko Krutwig-i Omenaldia. Iker* 10, 109–132. Bilbao: Euskaltzaindia.

Gorrotxategi, Joaquín. 1984. *Onomástica indígena de Aquitania*. Bilbao: University of the Basque Country.

Gorrotxategi, Joaquín. 1995. The Basque Language and its Neighbors in Antiquity. In José Ignacio Hualde, Joseba A. Lakarra and Robert Lawrence Trask (eds.), *Towards a history of the Basque language,* 31–63. Amsterdam: John Benjamins.

Greenberg, Joseph. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Joseph Greenberg (ed.), *Universals of Language*, 73–113. Cambridge, MA: MIT Press.

Haegeman, Liliane. 2006a. Conditionals, factives and the left periphery. *Lingua* 116. 1651–1669.

Haegeman, Liliane. 2006b. Argument fronting in English, Romance CLLD and the Left Periphery. In Raffaela Zanuttini, Héctor Campos, Elena Herburger, Paul Portner (eds.), *Negation, Tense and Clausal Architecture: Cross-Linguistic Investigations*, 27–52. Washington, D.C.: Georgetown University Press.

Heycock, Caroline. 2006. Embedded root phenomena. In Martin Everaert, Henk van Riemsdijk, (eds.), *The Blackwell Companion to Syntax*, vol. 2, 174–209. Oxford: Blackwell.

Hualde, José Ignacio. 2003. Introduction. In José Ignacio Hualde and Jon Ortiz de Urbina (eds.), *A Grammar of Basque*, 1–14. Berlin: Mouton de Gruyter.

Hualde, José Ignacio and Jon Ortiz de Urbina (eds.), *A Grammar of Basque*. Berlin: Mouton de Gruyter.

Kempchinsky, Paula. 2009. What can the subjunctive disjoint reference effect tell us about the subjunctive. *Lingua* 119. 1788–1810.

Kiparsky, Paul and Carol Kiparsky. 1970. Fact. In Danny D. Steinberg and Leon A. Jakobovits (eds.), *Semantics. An interdisciplinary reader in philosophy, linguistics and psychology*, 345–369. Cambridge: Cambridge University Press.

Lafitte, Piarres. 1979 [1944]. *Grammaire Basque (Navarro-Labourdin Littéraire)* (Bayonne, Librairie "Le Livre"). Baiona: Elkar.

Lafon, René. 1980 [1943]. *Le systéme du verbe basque au XVI^{e} siècle* (Bordeaux: Delmas, Publications de l'Université de Bordeaux). San Sebastian: Elkar

Laka, Itziar. 1990. *On the Nature of Functional Categories and Projections*. Cambridge, MA: MIT PhD dissertation.

Laka, Itziar. 1991. Sentence negation in Basque. In Joseba A. Lakarra (ed.), *Memoriae L. Mitxelena Magistri Sacrum*, 899–926. San Sebastian: Supplements of ASJU.

Laka, Itziar. 1993. The Structure of Inflection: A Case Study in X^0 Syntax. In José Ignacio Hualde and Jon Ortiz de Urbina (eds.) *Generative Studies in Basque Linguistics*, 21–70 Amsterdam & Philadelphia: John Benjamins.

Laka, Itziar. 2005. Deriving Split ergativity in the progressive: the case of Basque. In Alana Johns, Diane Massam and Juvenal Ndayiragije (eds.) *Ergativity: Emerging Issues*, 173–195. Dordrecht: Springer.

Laka, Itziar. 2006. On the nature of case in Basque: structural or inherent?. In Hans Broekhuis, Norbert Corver, Jan Koster, Riny Huybregts & Ursula Kleinhenz (eds.), *Organizing Grammar*, 374–382. Berlin: Mouton de Gruyter.

McCloskey, James. 2005. Questions and questioning in a local English. In Raffaela Zanuttini, Héctor Campos, Elena Herburger and Paul H. Portner (eds.), *Cross-Linguistic research in syntax and semantics: Negation, tense and clausal architecture*, 87–126. Washington: Georgetown University Press.

Mitxelena, Koldo. 1954. De onomástica aquitana. *Pirineos* 10. 409–458. [Reprinted in Mitxelena 1985, *Lengua e historia.* Madrid: Paraninfo. 409–445].

Mitxelena, Koldo. 1977. *Fonética Histórica Vasca*, 2nd edition. San Sebastián: Diputación Foral de Guipúzcoa.

Mitxelena, Koldo. 1985. *Lengua e historia.* Madrid: Paraninfo.

Nafarroako Gobernua. 2009. *Euskararen egoera Nafarroan 2008.* [The situation of the Basque language in Navarre 2008]. Iruñea [Pamplona]: Nafarroako Gobernua.

Nichols, Lynn. 2001. *The syntax and morphosyntax of propositional attitude complementation*. Cambridge, MA: Ms. Harvard University

Ormaechevarria, Ignacio. 1959. *Euskera. Un poco de gramática y algo de morfología del verbo vasco*. Zarautz: Itxaropena.

Ortiz de Urbina, Jon. 1989a. *Parameters in the Grammar of Basque*. Dordrecht: Foris.
Ortiz de Urbina, Jon. 1989b. Dislocaciones verbales en estructuras de polaridad. *International Journal of Basque Linguistics and Philology* 23(2). 393–410.
Ortiz de Urbina, Jon. 1995. Residual Verb Second and Verb First in Basque. In Katalin. É. Kiss (ed.), *Discourse Configurational Languages*, 99–121. Oxford: Oxford University Press.
Ortiz de Urbina, Jon. 1999. Force Phrases, Focus Phrases and Left Heads in Basque. In Jon Franco & Alazne Landa (eds.), *Grammatical Analysis of Romance and Basque*, 179–194. Amsterdam: John Benjamins.
Oyharçabal, Beñat. 1988. Operatzaile isila euskarazko perpaus erlatiboetan [The null operator in Basque relative clauses]. *International Journal of Basque Linguistics and Philology* 22(1). 93–97.
Oyharçabal, Beñat. 1997. La situation de la langue basque en Pays Basque Nord. *Lapurdum* 2. 45–62.
Oyharçabal, Beñat. 2003. Relatives. In José. I. Hualde & Jon Ortiz de Urbina (eds.), *A Grammar of Basque*, 762–823. Berlin: Mouton de Gruyter.
Rijk, Rudolf P. G. de. 1969. Is Basque S.O.V.?. *Fontes Linguae Vasconum* 1. 319–351. [Reprinted in de Rijk 1998, *De Lingua Vasconum: Selected Writings*. Bilbao: Supplements of ASJU. 13–38].
Rijk, Rudolf P. G. de. 1972. Partitive Assignment in Basque. *International Journal of Basque Linguistics and Philology* 6. 130–173. [Reprinted in de Rijk 1998, *De Lingua Vasconum: Selected Writings*. Bilbao: Supplements of ASJU. 151–182]
Rijk, Rudolf P.G. de. 2008. *Standard Basque. A progressive grammar.* Cambridge, MA: MIT Press.
Rivero, María Luisa & Angela Ralli (eds.). 2001. *Comparative syntax of Balkan languages*. Oxford: Oxford University Press.
Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.
Rizzi, Luigi. 1996. Residual Verb Second and the Wh-Criterion. In Adriana Belleti and Luigi Rizzi (eds.), *Parameters and Functional Heads*, 63–90. Oxford: Oxford University Press.
Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Liliane Haegeman (ed.), *Elements of Grammar*, 281–337. Dordrecht: Kluwer.
Sáez, Luis. 1989. Borrado/Forma Lógica: Comparativas con complemento nulo en español y en euskera. *International Journal of Basque Linguistics and Philology* 23(3). 677–700.
Trask, Robert Lawrence. 1995. Origin and relatives of the Basque language: review of the evidence. In José Ignacio Hualde, Joseba A. Lakarra and Robert Lawrence Trask (eds.), *Towards a history of the Basque language,* 65–99. Amsterdam: John Benjamins.
Uribe-Etxebarria, Myriam. 1994. *Interface Licensing Conditions on Negative Polarity Items: A Theory of Polarity and Tense Interactions*. University of Connecticut PhD dissertation.
Wiklund, Anna-Lenna, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson and Þorbjörn Hróarsdóttir. 2009. On the distribution and illocution of V2 in Scandinavian *that*-clauses. *Lingua* 119. 1914–1938.
Zuazo, Koldo. 1998, Euskalkiak, gaur [The Basque dialects today]. *Fontes Linguae Vasconum* 78. 191–234.
Zuazo, Koldo. 2008. *Euskalkiak. Euskararen dialektoak* [The dialects of Basque]. Donostia: Elkar.

Albert Borg and Ray Fabri

Semantic functions of complementizers in Maltese

1 Introduction

In this paper, we discuss complementation in Maltese, focusing on complementizers and their semantics. We distinguish between subordinate clauses functioning as complements as opposed to their function as modifiers. The focus will be mainly on complement clauses introduced by the following complementizers (the glosses give the approximate meanings): *li* 'that', *jekk* 'if' and *biex* '(in order) to'. The aim of this paper is to describe and discuss the distribution of these complementizers with different types of predicates in order to arrive at a characterisation of the meaning associated with each.

In Section 2, we place Maltese very briefly in its historical context. In Section 3, we briefly discuss some relevant formal (syntactic, morpho-phonological) features of Maltese, as well as its orthography. In Section 4, we introduce and discuss the three complementizers *li, jekk* and *biex*; we also consider the lack of an explicit complementizer. In Section 5, we contrast subordinate clauses which function as complements with those which function as modifiers. Section 6 focuses on the distribution of the complementizers. We conclude with Section 7.

2 Historical background

In origin, Maltese is a North African Arabic dialect brought to the Maltese Islands by the Arabs around 870 A.D. Cut off from the Arabic speaking world in 1249 and through intense contact with Old Sicilian followed by Italian in later centuries, it developed into a language in its own right, with the Romance influence evident in its grammar and lexis. Over the last two centuries or so, the influence of English has become increasingly significant, and Modern Maltese borrows extensively from English, adapting loan words to native morphological patterns, to varying degrees. Maltese is currently one of the official languages of the EU (cf. Borg 2012, Brincat 2011, Fabri 2010 for a detailed exposition and references). There are over 400,000 speakers of Maltese living in the Maltese Islands and at least as many migrants, chiefly in Australia but also in the U.S., Canada and the U.K.

3 Structure

In this paper, we are using Standard Maltese orthography in the examples. The following Table shows the graphemes and their corresponding sound values.

Grapheme	*Phoneme/s*		*Grapheme*	*Phoneme/s*	
a	/ɐ/	/ɐː/	l	/l/	
b	/b/		m	/m/	
ċ	/tʃ/		n	/n/	
d	/d/		o	/ɔ/	/ɔː/
e	/ɛ/	/ɛː/	p	/p/	
f	/f/		q	/ʔ/	
ġ	/dʒ/		r	/r/	
g	/g/		s	/s/	
għ	generally silent		t	/t/	
h	generally silent		u	/ʊ/	/uː/
ħ	/h/		v	/v/	
i	/ɪ/	/iː/	w	/w/	
ie	/ɪː/		x	/ʃ/	/ʒ/
j	/j/		ż	/z/	
k	/k/		z	/ts/	
			zz	/dzː/	

3.1 Phonology and morphology

In comparison to Modern Standard Arabic, Maltese lacks a number of phonemes, mostly back consonants, such as /ɣ/, /ʕ/ and /q/, and emphatics, such as /t'/, /s'/. In contrast, some phonemes in Modern Maltese, such as /p/ and /g/, are of Romance origin (Brincat 2011, Aquilina 1959, Borg 1997). However, perhaps the most striking characteristic of Modern Maltese is its hybrid morphology, with root-based and stem-based forms (cf. Mifsud 1995, Fabri 2009 and Spagnol 2011).

3.1.1 The definite article

In order for the reader to follow the examples, we need to briefly describe the definite article in Maltese since it is subject to morpho-phonemic variation which is reflected in the orthography. The underlying form of the definite article, which is cliticized to the word it precedes, is /l/ (orthographically *l-*) (see Camilleri 2009; Fabri 1993, 1996). Note that, although nouns in Maltese are typically marked for

gender and number, in the examples below, we only refer to them in the glosses when relevant to the point being made.

(1) a. *arblu*
mast
'mast'
b. *l-arblu*
DEF-mast
'the mast'
c. *isem*
name
'name'
d. *l-isem*
DEF-name
'the name'

Depending on the phonological environment it occurs in, /l/ undergoes the phonological processes of /i/-epenthesis and consonant assimilation. There are two forms of /i/-epenthesis: 'outer' and 'inner' epenthesis. Outer epenthesis takes place when the word that /l/ is affixed to starts with a consonant and the article is not preceded by a word ending in a vowel.

(2) a. *il-grada*
DEF-gate
'the gate'
b. *Fta*ħ*-na* *l-grada.*
open-1PL.PFV DEF-gate
'We opened the gate.'
c. *Feta*ħ *il-grada.*
open.3M.SG.PFV DEF-gate
'He opened the gate.'

Inner epenthesis occurs when the noun starts with a consonant cluster having initial [s,ʃ] (graphemically *s* and *x*).

(3) a. *spiżjar*
pharmacist
'pharmacist'
b. *l-ispiżjar*
pharmacist
'the pharmacist'
c. *xkora*
sack
'sack'
d. *l-ixkora*
DEF-sack
'the sack'

If the word that the article is attached to begins with the coronal consonants /d/, /t/, /s/, /z/, /ʃ/, /ts/, /tʃ/, /n/, /r/ (graphemically *d, t, s, ż, x, z, ċ, n, r*), regressive assimilation takes place. Note that /dʒ/ (graphemically *ġ*) is the only coronal consonant that does not trigger assimilation as in (6) below.

(4) a. *dar*
house
'house'
b. *id-dar*
DEF-house
'the house'
c. **il-dar*
DEF-house
'the house'

(5) a. *zija*
aunt
'aunt'
b. *iz-zija*
DEF-aunt
'the aunt'
c. **il-zija*
DEF-aunt
'the aunt'

(6) a. *ġita*
outing
'outing'
b. *il-ġita*
DEF-outing
'the garden'
c. **iġ-ġita*
DEF-outing
'the outing'

3.2 Syntax

3.2.1 Constituent order

Maltese has a relatively free constituent order on sentence level (Fabri & Borg 2002), allowing both SV and VS for intransitive sentences and SVO, SOV, OSV, OVS and VOS for transitive sentences. VSO is excluded when the direct object is definite, but allowed when the complement is an indefinite noun phrase or a locative expression. The unmarked order is SVO but, given the appropriate pro-

sodic configuration (intonation pattern, pause, etc.), the other variants can be used for different communicative effects, for example to express contrastivity. If a pronominal clitic agreeing with the direct object is attached to the verb, the order VSO is also possible. The following are a few examples for illustration in which commas roughly indicate prosodic effects.

(7) a. *It-tifla* *ħasl-et* *il-kelb.*
DEF-girl wash-3F.SG.PFV DEF-dog
'The girl washed the dog.'
b. **Ħasl-et,* *it-tifla,* *il-kelb.*
wash-3F.SG.PFV DEF-girl DEF-dog
c. *It-tifla* *marr-et* *fi-l-ġnien.*
DEF-girl go-3F.SG.PFV in-DEF-garden
'The girl went into the garden.'
d. *Marr-et,* *it-tifla,* *fi-l-ġnien.*
go-3F.SG.PFV DEF-girl in-DEF-garden
'The girl went into the garden.'
e. *Ħasl-it-u,* *t-tifl-a,* *l-kelb.*
wash-3F.SG.PFV.SBJ-3M.SG.OBJ DEF-girl-F.SG DEF-dog.M.SG
'As for the dog, the girl washed it.'

3.2.2 Topic

Following Fabri & Borg (2002) and Fabri (2010), Maltese can be characterised as a discourse configurational (Kiss 1995) language, especially in its spoken form: most constituents of the sentence can be topicalized by being placed at the beginning of the sentence, and marked as such by the appropriate suprasegmental configuration (Borg & Azzopardi-Alexander 2009). The topic constitutes one tone group and is separated from the following main clause, which constitutes another tone group with an intonation pattern resembling that found with an unmarked sentence. (8) is an example of an unmarked construction.

(8) *Il-kelb* *attakka* *l-qattusa.*
DEF-dog attack.3M.SG.PFV DEF-cat
'The dog attacked the cat.'

In (9), the direct object *il-qattusa* 'the cat' is a topic. Thus, it constitutes a separate tone group from the main clause and carries an intonation pattern signalling it is marked.

(9) *Il-qattus-a,* *il-kelb* *attakka-ha.*
DEF-cat-F.SG DEF-dog.M.SG attack.3M.SG.PFV.SBJ-3F.SG.OBJ
'As for the cat, the dog attacked it.'

There is a potential (non-obligatory) short pause between the two tone groups, which, when realized, accentuates the topicalized expression further. Note that in (9) there is pronominal encliticisation on the verb, co-referential with the object topic.

3.2.3 Case marking

Maltese has differential case marking with *lil* as the only case marker (glossed as CSE in the examples). The occurrence of *lil* before both direct and indirect objects is sensitive to degrees of animacy and definiteness (Borg & Mifsud 2002). The case marker is subject to morpho-phonemic variation triggered by the surrounding phonological context. The following are relevant examples.

(10) a. *It-tifel* *ra* *l-film.*
DEF-boy saw.3M.SG.PFV DEF-film
'The boy saw the film.'

b. *Il-pulizija* *qabad* *lil* *Pawlu/sieħb-u.*
DEF-police catch.3M.SG.PFV CSE Paul/friend-3M.SG.POSS
'The policeman caught Paul/his friend.'

c. *It-tifel* *ra* *'l* *Pawlu/sieħb-u.*
DEF-boy saw.3M.SG.PFV CSE Paul/friend-3M.SG.POSS
'The boy saw Paul/his friend'

d. *Il-pulizija* *qabad* *l-ir-raġel.*
DEF-police catch.3M.SG.PFV CSE-DEF-man
'The policeman caught the man.'

(11) a. *Il-ħalliel* *bagħat* *ittra* *l-qorti.*
DEF-thief send.3M.SG.PFV letter DEF-court
'The thief sent a letter to court.'

b. *Il-ħalliel* *bagħat* *ittra* *lil-l-Qorti.*
DEF-thief send.3M.SG.PFV letter CSE-DEF-court
'The thief sent a letter to the Court.'

3.2.4 Clause linkage

Clauses can be syntactically linked either via coordination or subordination. There are two basic types of subordinate clauses, namely, complement and modifying clauses, with the latter being either adverbial clauses or relative (i.e., adjectival) clauses. (See Camilleri 2011 for relative clauses in Maltese.) We discuss the difference between complement and adverbial clauses in detail in Section 5.

4 Complementation

Complement clauses are normally arguments of the verb (cf. Noonan [1985] 2007, for this definition of complementation), as in (12).

(12) *Il-mara* *qal-et* [***li*** *t-tifla* *ħasl-et* *il-kelb*].
DEF-woman say-3F.SG.PFV **that** DEF-girl wash-3F.SG.PFV DEF-dog
'The woman said that the girl washed the dog.'

In (12) the complement clause is the object of the verb and is introduced by the complementizer *li* (glossed as 'that' throughout these examples). If *li* is removed, the clause *t-tifla ħaslet il-kelb* can stand on its own as a complete sentence, i.e., it can be an independent clause. Note that this does not mean that the removal of a complementizer in Maltese necessarily results in an independent clause (see Section 4.2.1).

Complement clauses can also appear as arguments to other parts of speech, namely, to nouns, adjectives, and adverbs, and, in each case, the finite complement is introduced by *li* (Borg 1994b).

Sentence (13) provides an example of a sentential complement to a noun.

(13) *L-idea* [***li*** *t-tifla* *ħasl-et* *il-kelb*] *hija* *waħda* *sorprendenti*.
DEF-idea **that** DEF-girl wash-3F.SG.PFV DEF-dog COP.3F.SG one surprising
'The idea that the girl washed the dog is a surprising one.'

Sentence (14) exemplifies a complement to a (predicative) adjective.

(14) *Il-mara* *kuntenta* [***li*** *t-tifla* *ħasl-et* *il-kelb*].
DEF-woman happy **that** DEF-girl wash-3F.SG.PFV DEF-dog
'The woman is happy that the girl washed the dog.'

(15) shows an example of a complement to an adverb.

(15) *Issa* [***li*** *t-tifla* *ħasl-et* *il-kelb*] *il-mara* *kuntenta*.
now **that** DEF-girl wash-3F.SG.PFV DEF-dog DEF-woman happy
'Now that the girl has washed the dog, the woman is happy.'

Sentential complements can be nominalized and, therefore, can occur as nominal complements, in which case *li* does not figure (Borg 1994a). Thus, (16) is the nominalized version of the transitive sentence (17).

(16) *Il-ħasil* *ta-l-kelb* *mi-t-tifla*
DEF-washing of-DEF-dog from-DEF-girl
'The washing of the dog by the girl.'

(17) *It-tifla* *ħasl-et* *il-kelb.*
DEF-girl wash-3F.SG.PFV DEF-dog
'The girl washed the dog.'

The nominalized sentence can then appear as, for instance, the subject argument to a verb, as in (18).

(18) *Il-ħasil* *ta-l-kelb* *mi-t-tifla* *għoġob* *lil-l-mara.*
DEF-washing of-DEF-dog from-DEF-girl please.3M.SG.PFV CSE-DEF-woman
'The washing of the dog by the girl pleased the woman.'

Note that (16) is a noun phrase, headed by the noun *ħasil*; it is not sentence-like and no complementizer figures. The present study is concerned with finite, sentence-like complements, and, therefore, no more will be said about nominalized complements here.

li can introduce both a complement clause (13 repeated here as 19a) and a relative clause (19b), as can be seen from the following examples.

(19) a. *L-idea* [***li*** *t-tifla* *ħasl-et* *il-kelb*] *hija* *waħda*
DEF-idea **that** DEF-girl wash-3F.SG.PFV DEF-dog COP.3F.SG one
sorprendenti.
surprising
'The idea that the girl washed the dog is a surprising one.'

b. *L-idea* [***li*** *ħareġ* *bi-ha* *Pawlu*] *hija* *waħda*
DEF-idea **that** come.out.3M.SG.PFV with-3F.SG Paul COP.3F.SG one
sorprendenti.
surprising
'The idea that Paul came up with is a surprising one.'

The difference in status between the subordinate sentences in (19) as complement or relative clause can be established on the basis of two properties. First, in a relative clause coreference with an antecedent outside of the clause itself is obligatory, but it is not obligatory in the case of a complement clause. Secondly, one can coordinate two subordinate clauses only if they are either both complement clauses or both relative clauses (or also adverbial clauses). The following is an example of the unacceptable coordination of a relative clause with a complement clause.

(20) **L-idea* [***li*** *ħareġ* *bi-ha* *Pawlu*] *u* [***li*** *t-tifla*
DEF-idea **that** come.out.3M.SG.PFV with-3F.SG Paul and that DEF-girl
ħasl-et *il-kelb*] *hija* *waħda* *sorprendenti.*
wash-3F.SG.PFV DEF-dog COP.3F.SG one surprising
'The idea that Paul came up with and that the girl washed the dog is a surprising one.'

4.1 The complementizer *li*

In view of the foregoing discussion, it is clear that *li* cannot be uniquely associated with complementation since it also introduces modifying (relative) clauses. For this reason, the term complementizer is applied to the expression *li* in this work whenever it functions as a complementizer. Otherwise, we use the term subordinator or subordinating conjunction.

Note that *li* is a shorter form of *illi*, which occurs less frequently both in speech and in writing. Etymologically, *li/illi* harks back to the Standard Arabic masculine singular relative pronoun *allaḏī*, which appears in many other dialects of Arabic in the reduced (and grammatically invariable) form *ellī*/*illī*.[1] The complementizer in Standard Arabic is *ʾanna* 'that'.

Sentence (12) above, reproduced here as (21), typifies what one could call a canonical case of complementation, with the complementizer *li* and a finite complement.

(21)	*Il-mara*	*qal-et*	[***li***	*t-tifla*	*ħasl-et*	*il-kelb*].
	DEF-woman	say-3F.SG.PFV	**that**	DEF-girl	wash-3F.SG.PFV	DEF-dog
	'The woman said that the girl washed the dog.'						

li occurs with a range of complement-taking predicates. In Section 6, we discuss the distribution of *li* and the other complementizers in detail.

4.2 Other complementizers

Apart from complements with *li*, in this study, we are also concerned with complements which are introduced by two other complementizers, namely, *jekk* 'whether' (Section 4.2.1) and *biex* 'to' (Section 4.2.2), as well as with complements which lack an explicit complementizer (sections 4.2.3 and 4.2.4). We also consider question words which can function as complementizers such as *meta* 'when' (Section 5.2).

4.2.1 The complementizer *jekk*

The etymology of *jekk* is unclear since there seems to be no close cognate in other Arabic dialects. Among the suggestions which have been put forward so far, the

1 We would like to thank our colleague, Manwel Mifsud, for the etymologies of the complementizers.

most plausible seems to be that originally advanced by Dessoulavy (1938: 51), according to which it derives from the middle syllable of the Arabic expression *('in) yak(un)* 'if it is/were'.

The complementizer *jekk* 'whether' occurs with a matrix predicate that expresses doubt or uncertainty with respect to the complement, such as *saqsa* 'ask' and *iddubita* 'doubt'.

(22) *Il-mara* *ddubita-t/saqsie-t* [***jekk*** *it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman doubt-3F.SG.PFV/ask-3F.SG.PFV **if** DEF-girl wash-3F.SG.PFV-UNC DEF-dog
'The woman doubted if/whether the girl had washed the dog.'

In combination with certain (though not all) complement-taking predicates in the negative, *jekk* forces an interpretation of uncertainty, as in the following examples.

(23) a. *Ryan* *ma* *ddeċidie-x* [***jekk*** *i-rid-x* *i-mur*].
Ryan NEG decide.3M.SG.PFV-NEG **if** 3M.SG.IPFV-want-UNC 3M.SG.IPFV-go
'Ryan hasn't decided whether he wants to go.'
b. *Ryan* *ma* *j-af-x* [***jekk*** *i-rid-x* *i-mur*].
Ryan NEG 3M.SG.PFV-know-NEG **if** 3M.SG.IPFV-want-UNC 3M.SG.IPFV-go
'Ryan doesn't know whether he wants to go.'

Not every negative complement-taking predicate can take a *jekk* complement as can be seen in (24). This means that the semantics of such complement-taking predicates does not allow for an uncertain interpretation when in the negative.

(24) **Ryan* *ma* *j-aħsib-x* [***jekk*** *i-rid-x* *i-mur*].
Ryan NEG 3M.SG.PFV-think-NEG **if** 3M.SG.IPFV-want-UNC 3M.SG.IPFV-go
*'Ryan doesn't think whether he wants to go'

In general, the suffix *-x* occurs with the pre-verbal particle *ma* as the canonical verb negator, as can be seen in (25a) (see also Comrie 1982, Borg & Azzopardi-Alexander 1997), and on its own in the imperative (25b).

(25) a. *It-tifla* *ma* *ħasl-it-x* *il-kelb.*
DEF-girl NEG wash-3F.SG.PFV-NEG DEF-dog
'The girl didn't wash the dog.'
b. *T-aħsil-x* *il-kelb.*
2SG.IMP-wash-NEG DEF-dog
'Don't wash the dog.'

In the *jekk* examples in (22) and (23), however, *-x* cannot be said to express negativity of the suffixed, embedded verb, as the glosses show. For example, the interpretation of (23) cannot be 'Ryan has not decided/does not know whether he does not want to go'. The proposition expressed by the subordinate clause is not being

negated, but is being questioned. Note that in the interlinear glosses, we use UNC (uncertainty) for *-x* in this context to distinguish it from NEG (negative) *-x* (25a, b).

Given the right context, one can have a genuinely negative predicate in the subordinate clause (as in [26a] compared to [26b]), in which case, presumably negative *–x* and uncertain *–x* coalesce (see also [28a] and [28b]).

(26) a. *Saqsie-ni* [***jekk*** *ma* *n-af-x* *x'* *inhi*
ask.3M.SG.PFV.SBJ-1SG.OBJ **if** NEG 1SG.IPFV-know-NEG/UNC what COP
r-risposta].
DEF-reply
'He asked me whether I didn't know what the answer was.'

b. *Saqsie-ni* [***jekk*** *n-af-x* *x'* *inhi* *r-risposta*].
ask.3M.SG.PFV.SBJ-1SG.OBJ **if** 1SG.IPFV-know-UNC what COP DEF-reply
'He asked me whether I knew what the answer was.'

However, for some speakers, it is possible not to obtain a negative interpretation when both *jekk* and *ma –x* occur in the complement clause, i.e., for such speakers, examples (26a) above and (26b) are equivalent in meaning, namely, 'He asked me whether I knew what the answer was.'

With the appropriate intonation, the complementizer *jekk* may be omitted as in (27), without there being any change in meaning.

(27) *Il-mara* *saqsie-t* [*it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman ask-3F.SG.PFV DEF-girl wash-3F.SG.PFV-UNC DEF-dog
'The woman asked whether the girl had washed the dog.'

Note that in its use as a complementizer in these examples, *jekk* always governs a clause whose verb has to bear the suffix *-x*. This means that if *jekk* is removed, what remains is not an acceptable independent clause. This is in contrast to what happens, on the one hand, with complements introduced by *li* (see Section 4.1) and, on the other hand, with cases in which *jekk* introduces an adverbial and cannot be omitted (see Section 5.1).

From the optionality of *jekk* and the obligatory presence of *-x*, it seems plausible to assume that *-x* takes on the function of *jekk* in marking the clause as a complement clause, expressing uncertainty. Interestingly, when the complement verb is fully negated with *ma* and *-x*, *jekk* cannot be omitted, as can be seen in (26a) and (26b) above, repeated here as (28a) and (28b).

(28) a. *Saqsie-ni* [***jekk*** *ma* *n-af-x* *x'* *inhi*
ask.3M.SG.PFV.SBJ-1SG.OBJ **if** NEG 1SG.IPFV-know-NEG/UNC what COP
r-risposta].
DEF-reply
'He asked me whether I didn't know what the answer was.'

b *Saqsie-ni [ma n-af-x x' inhi r-risposta].
ask.3M.SG.PFV.SBJ-1SG.OBJ NEG 1SG.IPFV-know-NEG/UNC what COP DEF-reply
*'He asked me I don't know what is the answer'

Finally, note that there are no restrictions on the tense/aspect of the complement verb in a clause with *jekk*. The following examples illustrate this point, with (29) having the embedded verb encoding habituality, and with (30) having future meaning.

(29) *Il-mara saqsie-t* [***(jekk)*** *it-tifla t-aħsil-x il-kelb*].
DEF-woman ask-3F.SG.PFV **(if)** DEF-girl 3F.SG.IPFV-wash-UNC DEF-dog
'The woman asked whether the girl washes the dog.'

(30) I*l-mara saqsie-t* [***(jekk)*** *it-tifla hi-x se t-aħsel*
DEF-woman ask-3F.SG.PFV **(if)** DEF-girl COP.3F.SG-UNC FUT 3F.SG.IPFV-wash
il-kelb].
DEF-dog
'The woman asked whether the girl was going to wash the dog.'

4.2.2 The complementizer *biex*

biex is attested in various Arabic vernaculars as a form similar to *bēš*, usually with the same meaning as in Maltese. It is usually explained as a compound word made up of the preposition *bi* (Arabic *bi* 'in, by, with') followed by *-iex*, which in turn is a fusion of Arabic *ayy šai'* 'which thing'. Note that *-iex* is suffixed to a small number of Maltese prepositions, as in *għaliex* 'why', *fiex* 'in what', *fuqiex* 'on what', *daqsiex* 'how large', *mniex* 'from what' and, occasionally, *t'iex* 'of what'.

The following is an example of a complement introduced by the complementizer *biex* 'in order to'.

(31) *Il-mara sfurza-t li-t-tifla* [***biex*** *t-aħsel il-kelb*].
DEF-woman force-3F.SG.PFV CSE-DEF-girl **to** 3F.SG.IPFV-wash DEF-dog
'The woman forced the girl to wash the dog.'

The proposition realized by the complement clause with *biex* expresses the outcome required by the 3-place matrix predicate s*forza* 'force'. This is in contrast to the adverbial subordinate clause introduced by *biex* as an adverbial conjunction expressing purpose, which will be discussed further in 5.2. *biex* in (31), as a complementizer, is optional as can be seen in (32).

(32) *Il-mara sfurza-t li-t-tifla* [*t-aħsel il-kelb*].
DEF-woman force-3F.SG.PFV CSE-DEF-girl 3F.SG.IPFV-wash DEF-dog
'The woman forced the girl to wash the dog.'

4.2.3 Subordinating conjunctions and complements

A range of complement clauses like the one in (33) occur without a complementizer (i.e., *li, jekk* and *biex)*, and instead are introduced by subordinating conjunctions which typically function as question words, as in (34). Examples of such conjunctions are *xi* 'what', *meta* 'when', *x'ħin* 'what time', *kif* 'how', *kemm* 'how much/many'. Note that these elements can also function as adverbial conjunctions (see Section 5.2).

(33) *Saqsie-t-ha* [*x'* *kien-et* *qed* *t-agħmel*].
ask-3F.SG.PFV.SBJ-3F.SG.OBJ what be-3F.SG.PFV PROG 3F.SG.IPFV-do
'She asked her what she was doing.'

(34) *X'* *kien-et* *qed* *t-agħmel?*
what be-3F.SG.PFV PROG 3F.SG.IPFV-do
'What was she doing?'

Similarly, *meta/x'ħin* in (35) introduces the complement to the verb *saqsa* 'ask'.

(35) *Il-mara* *saqsie-t* [*meta/x'* *ħin* *it-tifla* *ħasl-et* *il-kelb*].
DEF-woman ask-3F.SG.PFV when/what time DEF-girl wash-3F.SG.PFV DEF-dog
'The woman asked when/at what time did the girl wash the dog.'

The subordinate clause in (35) has the same syntactic status as the nominal argument *xi ħaġa* 'something' in (36).

(36) *Il-mara* *saqsie-t* [*xi ħaġa*].
DEF-woman ask-3F.SG.PFV something
'The woman asked something.'

However, with the appropriate intonational break between *saqsiet* and *meta/x'ħin*, it is possible to construe the subordinate clause in (35) as an adverbial clause of time (see Section 5.2 below), in the sense that the woman asked the question at the time the dog was washed: 'The woman asked (the question) at the moment when the girl washed the dog'.

The following is an example of the same construction with *kif* 'how'.

(37) *Il-mara* *saqsie-t* [*kif* *ħasl-et* *il-kelb*].
DEF-woman ask-3F.SG.PFV how wash-3F.SG.PFV DEF-dog
'The woman asked how she washed the dog.'

Here, too, given the appropriate intonation, the subordinate clause can be construed as an adverbial clause of time, with *kif* meaning 'when' instead of 'how'. In

this case, the meaning is 'The woman asked (the question) at the moment when she washed the dog'.

Kif can also introduce an adverbial clause of manner.

(38) *It-tifel* *pinġa* *l-istampa* [*kif* *immaġina-ha* *hu*].
DEF-boy draw.3SG.M.PFV DEF-picture how imagine.3M.SG.PFV.SBJ-3F.SG.OBJ he
'The boy drew the picture as he imagined it.'

The following is an example of a complement clause with *kemm* 'how much', which asks about the quantity involved, as in (39).

(39) *Il-tifel* *saqsa* [*kemm* *xtra-t* *pakketti* *t-tifla*].
DEF-boy ask.3M.SG.PFV how_much buy-3F.SG.PFV packets DEF-girl
'The boy asked how many packets the girl bought.'

The expression *kemm* can also have a function similar to *jekk* (see Section 4.2.1). As Borg (2011) points out, *kemm* can occur with the negative propositional attitude predicate *iddubita* 'doubt', among others, as in (40a).

(40) a. *N-iddubita* [*kemm* *mar* *il-Belt*].
1SG.IPFV-doubt how.much go.3M.SG.PFV DEF-city
'I doubt whether he went to Valletta.'
b. *N-iddubita* [*jekk* *mar-x* *il-Belt*].
1SG.IPFV-doubt if go.3M.SG.PFV-UNC DEF-city
'I doubt whether he went to Valletta.'

In this context, *kemm* does not strictly speaking express a physical quantity but rather quantifies the degree of probability that the complement subject did in fact go to Valletta or not and, using this sentence, the speaker gives his subjective opinion that he probably did not go. This use of *kemm* is to be distinguished from a purely quantificational one introducing an adverbial (see Section 4.2.3). The difference in meaning between the two sentences in (40) is very subtle. While the clause introduced by *jekk* leaves the question of whether the person involved actually went to Valletta open, the clause introduced by *kemm* questions whether the person actually went to Valletta, implying that he probably did not go. Note that, unlike *jekk*, *kemm* cannot be omitted.

kemm also occurs with the knowledge predicate *jaf*, however only in the negative.

(41) *Ma* *naf-x* [*kemm* *mar* *il-Belt*].
NEG know.1SG.IPFV-NEG how.much go.3M.SG.PFV DEF-city
'I don't know whether he went to Valletta.'

Once again, here, (41) with *kemm* is loaded more in favour of the non-occurrence of the event in the complement, while (42), with *jekk* is neutral between the occurrence or non-occurrence of the event.

(42) *Ma* *n-af-x* [***(jekk)*** *mar-x* *il-Belt*].
NEG 1SG-know-NEG **(if)** go.3M.SG.PFV-UNC DEF-city
'I don't know whether he went to Valletta.'

4.2.4 Verb sequences

A common construction in Maltese is one that consists of a series of inflected verbs which lack a complementizer, either obligatorily or optionally.

Thus in (43) we have the modal complement-taking predicate *tista'* 'can' taking a complement:

(43) a. *It-tifla* *t-ista'* [*t-aħsel* *il-kelb*].
DEF-girl 3F.SG.IPFV-can 3F.SG.IPFV-wash DEF-dog
'The girl can wash the dog.'
b. **It-tifla* *t-ista'* [***li/biex/jekk*** *t-aħsel* *il-kelb*].
DEF-girl 3F.SG.IPFV-can **that/to/if** 3F.SG.IPFV-wash DEF-dog
*'The girl can that/to/if wash the dog.'

Besides modal verbs like *tista'*, there are other classes of complement-taking predicates occurring in a construction without a complementizer. The following are examples including phasal predicates, such as *beda* 'begin, *kompla* 'continue', *reġa'* 'repeat', *waqaf* 'stop', e.g., (44a), positive achievement predicates, such as *pprova* 'try', *irnexxielu* 'succeed/manage', e.g., (44b), and acquisition of knowledge predicates, such as *nesa* 'forget', *ftakar* 'remember', e.g., (44c).

(44) a. *It-tifla* *bdie-t* [*t-aħsel* *il-kelb*].
DEF-girl begin-3F.SG.PFV 3F.SG.IPFV-wash DEF-dog
'The girl began to wash the dog.'
b. *It-tifla* *ppruva-t* [*t-aħsel* *il-kelb*].
DEF-girl try-3F.SG.PFV 3F.SG.IPFV-wash DEF-dog
'The girl tried to wash the dog.'
c. *It-tifla* *nsie-t* [*t-aħsel* *il-kelb*].
DEF-girl forget-3F.SG.PFV 3F.SG.IPFV-wash DEF-dog
'The girl forgot to wash the dog.'

Note that, in every case, the embedded verb must be in the imperfect form. Maltese lacks a specific infinitive form, and makes use of the imperfect form of the verb in places where some languages (e.g. Italian, English, German) typically make use of an infinitival form.

Another example of a verb sequence is that in (45).

(45) a. *Il-kantant* *ħareġ* [*j-għaġġel*].
DEF-singer go.out.3M.SG.PFV 3M.SG.IPFV-hurry
'The singer went out in a hurry.'
b. **Il-kantant* *ħareġ* [***li/biex/jekk*** *j-għaġġel*].
DEF-singer go.out.3M.SG.PFV **that/to/if** 3M.SG.IPFV-hurry
*'The singer went out that/to/if hurry'

In contrast to the previous examples, in which the embedded verb phrase is a complement, *jgħaġġel* 'hurry' in (45) can only be reasonably interpreted as an adverbial of manner. In Section 5.3, we discuss similar examples and probe the difference between complements and adverbial clauses of purpose.

It is possible to have sequences with more than two verbs, as can be seen in (46):

(46) *It-tifla* *t-rid* [*t-erġa'* *t-mur* *t-għaġġel*
DEF-girl 3F.SG.IPFV-want 3F.SG.IPFV-repeat 3F.SG.IPFV-go 3F.SG.IPFV-hurry
t-aħsel *il-kelb*].
3F.SG.IPFV-wash DEF-dog
'The girl wants to go again in a hurry to wash the dog.'

The relevant status of each individual predicate will be discussed in Section 5.3.

Note that, in all of the examples discussed up to now in this section, all the verbs in a sequence obligatorily share the same subject.

Another class of instances in which the complementizer cannot occur involves predicates which allow raising-to-object, including desiderative *xtaq* 'wish', *ried* 'want', and perception *ra* 'see', *sema'* 'hear/listen' (see also Section 6.6.5).

(47) a. *Il-mara* *xtaq-et* *li-t-tifla* [*t-aħsel* *il-kelb*].
DEF-woman wish-3F.SG.PFV CSE-DEF-girl 3F.SG.IPFV-wash DEF-dog
'The woman wished the girl would wash the dog.'
b. *Il-mara* *ra-t* *li-t-tifla* [*t-aħsel* *il-kelb*].
DEF-woman see-3F.SG.PFV CSE-DEF-girl 3F.SG.IPFV-wash DEF-dog
'The woman saw the girl wash the dog.'

With equi verbs like *ġiegħel* 'force', the complementizer is also absent (see also Section 6.6.5). In these constructions, since the understood subject of the embedded predicate is the same (equi) as the object (which is *lil* marked, see Section 3.2.3) of the matrix verb, the embedded verb appears to agree with the direct object of the matrix verb:

(48) *Il-mara* *ġiegħl-et* [*li-t-tifla* *t-aħsel* *il-kelb*].
DEF-woman force-3F.SG.PFV CSE-DEF-girl 3F.SG.IPFV-wash DEF-dog
'The woman made the girl wash the dog.'

In these cases, too, the embedded verb must be in the imperfect form.

We conclude that the lack of a complementizer discussed in this section indicates a closer relationship between the matrix predicate and the complement than obtains when the complementizer is present. The matrix predicate controls the complement predicate both in terms of referentiality and tense/aspect (see also Section 6.6.5).

5 Adverbial vs. complement clauses

In the discussion in sections 4.2.3 and 4.2.4, we point out the fact that one has to distinguish between the status of the subordinate clause as either a complement or a modifier. In this section, we focus on different types of adverbial clauses in order to distinguish them from complement clauses.

5.1 *jekk* with modifying clause

The distinction between subordinate clauses as complements or modifiers is found in subordinate clauses introduced by *jekk*. The following is an example of *jekk* as an adverbial conjunction introducing a conditional clause, which may or may not be fulfilled (*realis*) (Borg & Azzopardi-Alexander 1997).

(49) a. [***Jekk*** *it-tifla* *t-aħsel* *il-kelb*], *n-agħlaq* *il-bieb* *ta-l-ġnien.*
if DEF-girl 3F.SG.IPFV-wash DEF-dog 1SG.IPFV-close DEF-door of-DEF-garden
'If the girl washes the dog, I'll close the garden door.'

b. *Il-mara* *j-iddispjaċi-ha* [***jekk*** *it-tifla* *t-aħsel* *il-kelb*].
DEF-woman 3M.SG.IPFV.be.sorry-3F.SG **if** DEF-girl 3F.SG.IPFV-wash DEF-dog
'The woman will be sorry if the girl washes the dog.'

Note that although the subordinate clause is introduced by *jekk*, the verb here does not carry the negative suffix *-x*. This is a convenient way of distinguishing complement clauses from adverbial clauses when both are introduced by *jekk*. Compare the adverbial clauses in (49) to the complement clause in (50).

(50) *Il-mara* *saqsie-t* [***jekk*** *it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman ask-3F.SG.PFV **if** DEF-girl wash-3F.SG.PFV-UNC DEF-dog
'The woman asked if the girl had washed the dog.'

5.2 Other adverbial conjunctions

In Section 4.2.2 we discussed *biex* as complementizer. However, just like *jekk*, *biex* can also function as adverbial conjunction of purpose, as in (51).

(51) *Pawlu j-kanta arja* [***biex*** *j-ipprattika l-vuċi*].
Paul 3M.SG.IPFV-sing aria **to** 3M.SG.IPFV-practise DEF-voice
'Paul sings an aria to train his voice.'

In contrast to its use as complementizer (see [36] in 4.2.2), in its adverbial use, *biex* is obligatory (52).

(52) **Pawlu j-kanta arja* [*j-ipprattika l-vuċi*].
Paul 3M.SG.IPFV-sing aria 3M.SG.IPFV-practise DEF-voice
*'Paul sings an aria practises his voice.'

Other conjunctions introducing adverbial clauses are also possible, for instance *meta* 'when' in (53a and 53b), introducing an adverbial clause of time, and *għax* 'because' in (52c), introducing an adverbial clause of reason.

(53) a. *It-tifla ħasl-et il-kelb meta ra-t-u maħmuġ.*
DEF-girl wash-3F.SG.PFV DEF-dog when see-3F.SG.PFV.SBJ-3M.SG.OBJ dirty
'The girl washed the dog when she saw him dirty.'
b. *Il-mara t-iddejjaq meta t-tifla t-aħsel il-kelb.*
DEF-woman 3F.SG.IPFV-be.annoyed when DEF-girl 3F.SG.IPFV-wash DEF-dog
'The woman is annoyed when the girl washes the dog.'
c. Il-*mara iddejq-et għax it-tifla ħasl-et il-kelb.*
DEF-woman be.annoyed-3F.SG.PFV because DEF-girl wash-3F.SG.PFV DEF-dog
'The woman was annoyed because the girl washed the dog.'

A number of prepositions can occur with *li* introducing a clause, as shown in the following example:

(54) *It-tifla qal-et lil omm-ha wara li ħasl-et il-kelb.*
DEF-girl say-3F.SG.PFV CSE mother-3F.SG.POSS after that wash-3F.SG.PFV DEF-dog
'The girl told her mother after she washed the dog.'

The subordinate clause here functions as an adverbial modifier. In the case of (54) it is an adverbial clause of time, specifying when the girl told her mother.

Rather than a combination of a preposition and a sentential complement introduced by *li*, *wara li* can be treated as a case of a complex adverbial conjunction introducing an adverbial clause. This is obvious also from other phonologically integrated (lexicalized) forms such as *billi* 'since/in as much as', *malli* 'as soon as', *milli* 'from what', and *talli* 'because of', which are derived from a combination of a preposition, namely, *bi*, '(instrumental) with', *ma'* '(comitative) with',

minn 'from', *ta'* 'of', with *li*. In (55) we find *malli* introducing an adverbial time clause showing contemporaneity.

(55) *It-tifla* *ħasl-et* *il-kelb* *malli* *daħal* *mil-l-ġnien.*
DEF-girl wash-3F.SG.PFV DEF-dog as.soon.as enter.3M.SG.PFV from-DEF-garden
'The girl washed the dog as soon as it came in from the garden.'

5.3 Verb sequences as modifiers or complements

In Section 4.2.4, we mentioned constructions with verb sequences. An interesting question in this context is what the status of each predicate involved in the construction is, namely, whether it is a modifier or a complement. It is interesting to contrast (56a) with (56b).

(56) a. *Il-kantant* *ħareġ* [*j-għaġġel*].
DEF-singer go.out.3M.SG.PFV 3M.SG.IPFV-hurry
'The singer went out in a hurry.'
b. *Il-kantant* *ħareġ* [*j-ixtri*].
DEF-singer go.out.3M.SG.PFV 3M.SG.IPFV-buy
'The singer went out to buy something.'

These two constructions share two properties, namely, the two verbs must be coreferential and therefore agree with respect to person, gender and number, and the embedded verb must be in the imperfect form. In contrast, they differ in that the embedded predicate in (56a) modifies the action expressed by the matrix verb, whereas in (56b) the embedded verb expresses the purpose why the singer went out.

Although, on the surface, the two constructions appear to be structurally parallel, it can be shown that they are not. Thus, given a purpose interpretation, in (57b) but not in (57a) (see also [56a] and [56b]), we can have the conjunction *biex* separating the two predicates.

(57) a. **Il-kantant* *ħareġ* [***biex*** *j-għaġġel*].
DEF-singer go.out.3M.SG.PFV **to** 3M.SG.IPFV-hurry
*'The singer went out to hurry.'
b. *Il-kantant* *ħareġ* [***biex*** *j-ixtri*].
DEF-singer go.out.3M.SG.PFV **to** 3M.SG.IPFV-buy
'The singer went out to buy something.'

We assume that *jgħaġġel* in (56a) and (57a) is an adverbial modifier of manner while *jixtri* in (56b) and (57b) is a complement. The following are two other examples.

(58) a. *Sofia* *ħarġ-et* [*t-għajjat*].
Sophie go.out-3F.SG.PFV 3F.SG.IPFV-shout
'Sophie went out shouting.'
b. *Ray* *ġie* [*j-gerger*].
Ray come.3M.SG.PFV 3M.SG.IPFV-grumble
'Ray came grumbling.'

On a manner interpretation of (58a) and (58b), the question of the optionality of *biex* does not arise, since it would force a purpose reading. Note that often ambiguity in interpretation can arise between a manner (and modifier) reading and a purpose (and complement) reading when *biex* is absent, and when the second predicate allows it. For example, (58a) can be interpreted as 'Sophie went out to shout', in which case *biex* is possible but not obligatory. This is not the case with (56a) because the manner predicate *għaġġel* does not readily allow a purpose interpretation. We will not pursue the matter further here.

Without going into a detailed analysis at this point, we repeat here the example (46) given in Section 4.2.4, with multiple predicates in sequence.

(59) *It-tifla* *t-rid* [*t-erġa'* [[*t-mur* *t-għaġġel*]
DEF-girl 3F.SG.IPFV-want 3F.SG.IPFV-repeat 3F.SG.IPFV-go 3F.SG.IPFV-hurry
t-aħsel *il-kelb*]].
3F.SG.IPFV-wash DEF-dog
'The girl wants to go again in a hurry to wash the dog.'

On the basis of what we claim in the previous discussion, in (59) *terġa' tmur tgħaġġel taħsel il-kelb* is a complement of *trid*, *tmur tgħaġġel taħsel il-kelb* is a complement of phasal *terġa'*, *tgħaġġel* is a manner modfier of *tmur* and *taħsel il-kelb* is a complement of *tmur*. Note that we can optionally insert *biex* between *tgħaġġel* and *taħsel* as in (60).

(60) *It-tifla* *t-rid* [*t-erġa'* [[*t-mur* *t-għaġġel*] ***biex***
DEF-girl 3F.SG.IPFV-want 3F.SG.IPFV-repeat 3F.SG.IPFV-go 3F.SG.IPFV-hurry **to**
t-aħsel *il-kelb*]].
3F.SG.IPFV-wash DEF-dog
'The girl wants to go again in a hurry to wash the dog.'

6 The distribution of the complementizers

In this section, we focus on the distribution of the complementizers in relation to various types of complement-taking predicates in order to throw light on the semantics of the complementizers. We will examine complement-taking predicates with respect to whether they take only *li* (Section 6.1), only *biex* (Section 6.2)

or only *jekk* (Section 6.3), whether they allow a choice between more than one complementizer (sections 6.4 and 6.5), or whether they occur without a complementizer (Section 6.6). It turns out that factors like negation, tense/aspect and person, among others, both with respect to the matrix predicate and the complement verb, may result in different choices of complementizer. The following account is not meant to be exhaustive but simply gives an indication of the possible directions that future work might take. To keep things simple, we focus on examples with the subject of the matrix predicate in the 3rd person.

Also note that, in the following, we use the term *factive* to mean that the proposition encoded by the complement is *asserted* and *realized*. The distinction between asserted and realized is necessary in order to deal with cases such as complements in the future in which the proposition encoded by the complement is asserted but not realized. This difference will be relevant later on when we attempt to characterize the complementizers.

6.1 Only *li* but not *biex/jekk*

The following is an example with the (positive) propositional attitude predicate *ħaseb* 'think', which takes *li* but not *biex, jekk*.

(61) a. *Il-mara* *t-aħseb* [***li*** *t-tifla* *ħasl-et* *il-kelb*].
DEF-woman 3F.SG.IPFV-think **that** DEF-girl wash-3F.SG.PFV DEF-dog
'The woman thinks that the girl washed the dog.'

b. **Il-mara* *t-aħseb* [***jekk/biex*** *it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman 3F.SG.IPFV-think **if/to** DEF-girl wash-3F.SG.PFV-UNC DEF-dog
*'The woman thinks if/to the girl washed the dog.'

This is also the case when the matrix predicate is negative as in (62). Note that both the positive (61a) and negative (62a) sentences are non-factive.

(62) a. *Il-mara* *ma* *t-aħsib-x* [***li*** *t-tifla* *ħasl-et* *il-kelb*].
DEF-woman NEG 3F.SG.IPFV-think-NEG **that** DEF-girl wash-3F.SG.PFV DEF-dog
'The woman doesn't think that the girl washed the dog.'

b. **Il-mara* *ma* *t-aħsib-x* [***jekk/biex*** *it-tifla*
DEF-woman NEG 3F.SG.IPFV-think-NEG **if/to** DEF-girl
ħasl-it-x *il-kelb*].
wash-3F.SG.PFV-UNC DEF-dog
*'The woman doesn't think if/to the girl washed the dog.'

We note a similar distribution with the following predicates, among others: commentative (emotional) *iddispjaċih* 'be sorry/regret' and *iddejjaq* 'be annoyed', pretence *stħajjel* 'imagine' and *immaġina* 'imagine', positive/negative proposi-

tional attitude *assuma* 'assume', *qabel* 'agree', *emmen* 'believe' and *innega* 'deny, the predicate of fearing *inkwieta* 'worry', desiderative *ittama* 'hope', knowledge/acquisition of knowledge *irrealizza* 'realize' and *ħolom* 'dream', utterance *wiegħed* 'promise'. In the case of *iddispjaċih* 'be sorry/regret' and *iddejjaq* 'be annoyed', *inkwieta, qabel, irrealizza*, in both positive and negative, the complement is factive while for the rest it is non-factive.

The occurrence of *li* with a number of predicates, such as commentative (emotional) *iddejjaq* 'be annoyed', is sensitive to the aspectual nature of the embedded clause. *li* is obligatory if the complement clause expresses a state (63).

(63) a. *Pawlu j-iddejjaq* [***li*** *mhux irqiq*].
Paul 3M.SG.IPFV-be.annoyed **that** not thin.M.SG
'Paul is annoyed that he is not slim.'
b. **Pawlu j-iddejjaq* [*mhux irqiq*].
Paul 3M.SG.IPFV-be.annoyed not thin
*'Paul is annoyed not thin.'
c. *Pawlu j-iddejjaq* [***li*** *Marija mhux irqiq-a*].
Paul 3M.SG.IPFV-be.annoyed **that** Mary not thin-F.SG
'Paul is annoyed that Mary is not slim.'
d. **Pawlu j-iddejjaq* [*Marija mhux irqiq-a*].
Paul 3M.SG.IPFV-be.annoyed Mary not thin-F.SG
'Paul is annoyed Mary is not slim.'

When the complement expresses an activity, if the matrix and complement predicates are coreferential, *li* is not possible (64a) and (64b), while if they are not coreferential, *li* is obligatory (64c) and (64d).

(64) a. **Pawlu*$_i$ *j-iddejjaq* [***li*** *j-pejjep*$_i$ *ġewwa*].
Paul 3M.SG.IPFV-be.annoyed **that** 3M.SG.IPF-smoke inside
'Paul is annoyed that he smokes inside.'
b. *Pawlu*$_i$ *j-iddejjaq* [*i-pejjep*$_i$ *ġewwa*].
Paul 3M.SG.IPFV-be.annoyed 3M.SG.IPFV-smoke inside
'Paul doesn't like smoking indoors.'
c. *Pawlu jiddejjaq* [***li*** *Marija t-pejjep ġewwa*].
Paul 3M.SG.IPFV-be.annoyed **that** Mary 3F.SG.IPFV-smoke inside
'Paul doesn't like Mary to smoke indoors.'
d. **Pawlu j-iddejjaq* [*Marija t-pejjep ġewwa*].
Paul 3M.SG.IPFV-be.annoyed Mary 3F.SG.IPFV-smoke inside
'Paul is annoyed Mary smokes inside.'

Another predicate like *iddejjaq* is the predicate of fearing *beża'* 'be afraid'.

Some of the predicates discussed above allow optional *li* in every context (state/activity, coreferentiality, tense/aspect). (65) shows a few examples with *ħaseb*, involving tense/aspect in (65a), coreferentiality in (65b) and state/activity in (65c) and (65d).

(65) a. *Il-mara* *t-aħseb/ħasb-et* [***(li)*** *t-tifla*
DEF-woman 3F.SG.IPFV.think/think-3F.SG.PFV **that** DEF-girl
t-aħsel/ħasl-et *il-kelb*].
3F.SG.IPFV-wash/wash-3F.SG. PFV DEF-dog
'The woman thinks/thought that the girl washes/washed the dog.'

b. *Jien* *n-aħseb* [***(li)*** *n-aħsel/ħsil-t* *il-kelb*].
I 1SG.IPFV-think **that** 1SG.IPFV-wash/wash-1SG.IPFV DEF-dog
'I think that I (shall) wash/washed the dog.'

c. *Jien* *n-aħseb* [***(li)*** *Pawlu* *għand-u* *raġun*].
I 1SG.IPFV.think **that** Paul have-3M.SG reason
'I think that Paul is right.'

d. *Jien* *n-aħseb* [***(li)*** *Pawlu* *sejjer* *id-dar*].
I 1SG.IPFV.think **that** Paul go.3M.SG.PROG DEF-house
'I think that Paul is going home.'

6.2 *biex*

The predicate *sforza* 'force' can occur with (optional) *biex* but not *li* or *jekk*.

(66) a. *Pawlu* *sforza* *li-t-tifla* [***(biex)*** *t-aħsel* *il-kelb*].
Paul 3M.SG.PFV.force CSE-DEF-girl **to** 3F.SG.IPFV-wash DEF-dog
'Paul made the girl wash the dog.'

b. **Pawlu* *sforza* *li-t-tifla* [***li/jekk*** *t-aħsel* *il-kelb*].
Paul 3M.SG.PFV.force CSE-DEF-girl **that/if** 3F.SG.IPFV-wash DEF-dog
'Paul forced the girl that/if she washes the dog.'

Another example is the manipulative predicate *ordna* 'order', which also occurs with optional *biex* as in (67), in an equi construction.

(67) a. *Pawlu* *ordna* *li-t-tifla* [***(biex)*** *t-aħsel* *il-kelb*].
Paul 3M.SG.PFV.order CSE-DEF-girl **to** 3F.SG.IPFV-wash DEF-dog
'Paul ordered the girl to wash the dog.'

b. **Pawlu* *ordna* *li-t-tifla* [***li/jekk*** *t-aħsel* *il-kelb*].
Paul 3M.SG.PFV.order CSE-DEF-girl **that/if** 3FS.G.IPFV-wash DEF-dog
*'Paul ordered the girl that/if she washes the dog.'

However, when *ordna* occurs in the non-equi version and is intransitive, *li* is possible.

(68) *Il-Kummissarju* *ordna* [**li**) *l-pulizija* *j-ilbsu*
DEF-Commissioner order.3M.SG.PFV **that** DEF-police 3PL.IPFV-dress
l-uniformi *ta-x-xitwa*].
DEF-uniform of-DEF-winter
'The Commissioner ordered that the police put on their winter uniform.'

When *ordna* has a direct object noun phrase which is not corefential with the embedded subject noun phrase, *li* is impossible (69).

(69) *Il-mara ordna-t li-t-tifla [Marija t-aħsel il-kelb].
DEF-woman order-3F.SG.PFV CSE-DEF-girl Mary 3F.SG.IPFV-wash DEF-dog
*'The woman ordered the girl Mary washes the dog.'

Other similar predicates are *talab* 'demand' and manipulative *ikkmanda* 'order'.

The complement verb in these contexts always occurs in the imperfect form, which is to be expected, since the complement expresses a potential event or state (see Noonan 2007: 104), i.e., one that is expected to be realized.

(70) *Il-mara ordna-t li-t-tifla [**biex** ħasl-et il-kelb].
DEF-woman order-3F.SG.PFV CSE-DEF-girl **to** wash-3F.SG.PFV DEF-dog
*'The woman ordered the girl to washed the dog.'

Note that the matrix predicate in this case is always transitive, i.e. it must have a direct object.

(71) *Il-mara ordna-t [**biex** t-aħsel il-kelb].
DEF-woman order-3F.SG.PFV **to** 3F.SG.IPFV-wash DEF-dog
*'The woman ordered to wash the dog.'

When the predicate *qal* 'say' occurs with *biex*, it is no longer interpreted as an utterance predicate (see Section 6.5) but as manipulative.

(72) Il-mara qal-et li-t-tifla [**biex** t-aħsel il-kelb].
DEF-woman tell-3F.SG.PFV CSE-DEF-girl **to** 3F.SG.IPFV-wash DEF-dog
'The woman told the girl to wash the dog.'

The occurrence of the complementizer *biex* or *li* with *ikkonvinċa* and *ipperswada* depends on whether the subordinate clause encodes a non-stative (or active) situation (73a) or a stative (non-active) situation (73b), respectively

(73) a. Pawlu ipperswada li-t-tifla [**(biex)** t-aħsel il-kelb].
Paul persuade.3M.SG.PFV CSE-DEF-girl **to** 3F.SG.IPFV-wash DEF-dog
'Paul persuaded the girl to wash the dog.'
b. Pawlu ipperswada li-t-tifla [**li** għand-u ħafna flus].
Paul persuade.3M.SG.PFV CSE-DEF-girl **that** have-3M.SG much money
'Paul persuaded the girl that he has a lot of money.'

The intransitive predicates *ħaseb* and *ftakar* can also take a *biex* complement, in which case the meaning is that of something being planned or made provision for, as opposed to the meaning of a positive propositional attitude predicate, e.g., *ħaseb* 'think', or (acquisition of) knowledge predicate e.g., *ftakar* 'remember', when they occur with *li* (see [65] in Section 6.1).

(74) a. *Il-mara* *ħasb-et* [***biex*** *t-aħsel* *il-kelb*].
DEF-woman think-3F.SG.PFV **to** 3F.SG.IPFV.wash DEF-dog
'The woman made provisions to wash the dog.'
b. *Il-mara* *ftakr-et* [***biex*** *t-aħsel* *il-kelb*].
DEF-woman remember-3F.SG.PFV **to** 3F.SG.IPFV-wash DEF-dog
'The woman remebered to wash the dog.'

With verbs like *stinka* 'try hard' and *ipprepara* 'prepare', which cannot occur with *li* and *jekk*, *biex* is obligatory.

(75) *L-istudent* *stinka/ipprepara* [***biex***/***li**/***jekk*** *j-għaddi*
DEF-student try.hard/prepare.3M.SG.PFV ***to*/that/if** 3M.SG.IPFV-pass
mi-ll-eżami].
from-DEF-exam
'The student tried hard/prepared to pass the exam.'

An interesting case of the occurrence of *biex* is the following:

(76) *It-tifla* *bdie-t* [***biex*** *t-aħsel* *il-kelb*].
DEF-girl begin-3F.SG.PFV **to** 3F.SG.IPFV-wash DEF-dog
'The girl started by washing the dog.'

Here, *biex* does not function as an adverbial conjunction with the meaning 'in order to (purpose)' but forms part of a compound predicate [*beda* + *biex*], which means 'start by'. The pattern [V + *biex*] can be found in other examples, such as *kompla biex* 'continued by', *wasal biex* 'came to the point of', *spiċċa biex* 'ended up by'. We will not discuss this any further here.

6.3 Only *jekk* but not *li*

The utterance predicate *saqsa* 'he asked' seems to be the only predicate that can occur with the complementizer *jekk* to the exclusion of *li* and *biex*.

Please refer to the discussion of this predicate in the earlier Section 4.2.1 on the complementizer *jekk*. In particular, note sentences (77a) and (77b) with a positive and a negative matrix verb, respectively.

(77) a. *Il-mara* *saqsie-t* [***jekk*** **/*li*** *it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman ask-3F.SG.PFV **if/that** DEF-girl wash-3F.SG.PFV-UNC DEF-dog
'The woman asked if the girl had washed the dog.'
b. *Il-mara* *ma* *saqsie-t-x* [***jekk***/***li*** *it-tifla* *ħasl-it-x*
DEF-woman NEG ask-3F.SG.PFV-NEG **if/that** DEF-girl wash-3F.SG.PFV-UNC
il-kelb].
DEF-dog
'The woman didn't ask whether the girl had washed the dog.'

As example (78) shows, *jekk* may be omitted, in which case the complement verb must still carry the suffix *-x*.

(78) *Il-mara* *saqsie-t* [*it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman ask-3F.SG.PFV DEF-girl wash-3F.SG.PFV-UNC DEF-dog
'The woman asked whether the girl had washed the dog.'

In the written form, the absence of *jekk*, as in in (78), results in ambiguity, as shown in (79), between (i) with intransitive *saqsa*, on the one hand, and (ii), with transitive *saqsa*, on the other.

(79) *Il-mara* *saqsie-t* [*it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman ask-3F.SG.PFV DEF-girl wash-3F.SG.PFV-UNC DEF-dog
i. 'The woman asked whether the girl had washed the dog.'
ii. 'The woman$_i$ asked the girl$_j$ whether she$_j$ had washed the dog.'

In the spoken form, this ambiguity is resolved through the appropriate intonation, while, in both the written and spoken forms, the ambiguity disappears, once the object marker *lil* (assimilated to *lit-* in this case) is made explicit, as in (80).

(80) *Il-mara* *saqsie-t* *li-t-tifla* [*ħasl-it-x* *il-kelb*].
DEF-woman ask-3F.SG.PFV CSE-DEF-girl wash-3F.SG.PFV-UNC DEF-dog
'The woman asked the girl whether she had washed the dog.'

Of course, in (80) there is no ambiguity with respect to a transitive or intransitive reading of *saqsa*, but there is ambiguity with respect to the pronominal binding of 'she' with 'woman', 'girl' or some other referent in context; we will not pursue this further here.

6.4 Both *jekk* and *li*

Nesa 'forget' is an example of a predicate that can take either *li* or *jekk*, whether it is positive or negative.

(81) a. *Il-mara* *insie-t* [***li*** *t-tifla* *ħasl-et* *il-kelb*].
DEF-woman forget-3F.SG.PFV **that** DEF-girl wash-3F.SG.PFV DEF-dog
'The woman forgot that the girl washed the dog.'
b. *Il-mara* *ma* *nsie-t-x* [***li*** *t-tifla* *ħasl-et* *il-kelb*].
DEF-woman NEG forget-3F.SG.PFV-NEG **that** DEF-girl wash-3F.SG.PFV DEF-dog
'The woman did not forget that the girl washed the dog.'
c. *Il-mara* *insie-t* [***jekk*** *it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman forget-3F.SG.PFV **if** DEF-girl wash-3F.SG.PFV-UNC DEF-dog
'The woman forgot whether the girl washed the dog.'

d. *Il-mara* *ma* *nsie-t-x* [***jekk*** *it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman NEG forget-3F.SG.PFV-NEG **if** DEF-girl wash-3F.SG.PFV-UNC DEF-dog
'The woman did not forget whether the girl had washed the dog.'

Note that in (81a) and (81b) *li* occurs with a factive complement, while with *jekk* in (81c) and (81d), the complement is non-factive.

6.5 *li* throughout and *jekk* in certain contexts

The positive propositional attitude predicate *ammetta* 'admit' occurs with *li* in both positive and negative contexts.

(82) a. *Il-mara* *ammettie-t* [***li*** *t-tifla* *ħasl-et* *il-kelb*].
DEF-woman admit-3F.SG.PFV **that** DEF-girl wash-3F.SG.PFV DEF-dog
'The woman admitted that the girl washed the dog.'
b. *Il-mara* *ma* *ammettie-t-x* [***li*** *t-tifla* *ħasl-et* *il-kelb*].
DEF-woman NEG admit-3F.SG.PFV-NEG **that** DEF-girl wash-3F.SG.PFV DEF-dog
'The woman didn't admit that the girl washed the dog.'

In a negative context *ammetta* also takes *jekk*, both in a declarative (83a) and an interrogative (83b) context.

(83) a. *Il-mara* *ma* *ammettie-t-x* [***jekk*** *it-tifla* *ħasl-it-x*
DEF-woman NEG admit-3F.SG.PFV-NEG **if** DEF-girl wash-3F.SG.PFV-UNC
il-kelb].
DEF-dog
'The woman didn't admit whether the girl washed the dog.'
b. *Il-mara* *ma* *ammettie-t-x* [***jekk*** *it-tifla* *ħasl-it-x*
DEF-woman NEG admit-3F.SG.PFV-NEG **if** DEF-girl wash-3F.SG.PFV-UNC
il-kelb]?
DEF-dog
'Didn't the woman admit whether the girl washed the dog?'

When positive, it can also take *jekk* if the context is interrogative.

(84) a. *Il-mara* *ammettie-t* [***jekk*** *it-tifla* *ħasl-it-x* *il-kelb*]*?*
DEF-woman admit-3F.SG.PFV **if** DEF-girl wash-3F.SG.PFV-UNC DEF-dog
'Did the woman admit whether the girl washed the dog?'
b. **Il-mara* *ammettie-t* [***jekk*** *it-tifla* *ħasl-it-x* *il-kelb*].
DEF-woman admit-3F.SG.PFV **if** DEF-girl wash-3F.SG.PFV-UNC DEF-dog
*'The woman admitted if the girl washed the dog.'

Predicates that behave in the same way are *ftakar* 'remember' (predicate of knowledge/acquisition of knowledge) and *ra* 'see' (perception), *qal* 'say', *iddik-*

jara 'declare' (utterance predicates) and *innega* 'deny' (negative propositional attitude).

The above discussion indicates that *jekk* gives rise to an uncertain interpretation of the complement, and that it can occur either in combination with a matrix predicate that expresses uncertainty lexically, such as *saqsa* 'ask' (see Section 4.2.1), or with matrix predicates, such as *ammetta 'admit'*, which, when negated or questioned, allow an uncertain reading of the complement.

6.6 Obligatory lack of a complementizer

A number of complement-taking predicates take complements obligatorily with no overt complementizer. We will not go into whether these are better analysed as paratactic constructions or serial verb constructions or whether they should be analysed as involving a "zero-complementizer". (See Noonan 2007 for some discussion and references, Maas 2007 for a consideration of complex predicates in Maltese and Arabic, and Stolz 2009 for verb chains in modern Maltese.)

An example of a predicate which requires a complement clause obligatorily without a complementizer is the equi predicate *pprova* 'try'.

(85) a. *Pawlu pprova [j-aħsel il-kelb].*
Paul try.3M.SG.PFV 3M.SG.IPFV-wash DEF-dog
'Paul tried to wash the dog.'

b. **Pawlu pprova* [***biex*/*li*/*jekk*** *j-aħsel il-kelb*].
Paul try.3M.SG.PFV **to**/**that**/**if** 3M.SG.IPFV-wash DEF-dog
*'Paul tried to/that/if washes the dog.'

This class of predicates includes the predicates discussed in the following sections. The predicates are classified on the basis of both semantic (modality, phasality) and morpho-syntactic (pseudo-predicates, etc.) criteria, and, therefore, some predicates can belong to more than one class, such as *għandu* 'have to', which is a modal pseudo-predicate.

6.6.1 Modal verbs

Examples of modal verbs are *seta'* 'can' and *ried* 'has to'.

(86) a. *Pawlu j-ista' j-oqgħod għal-l-eżami.*
Paul 3M.SG.IPFV-can 3M.SG.IPFV-stay for-DEF-exam
'Paul can take/stay for the exam.'

b. *Marija t-rid* [*t-oħroġ minnufih*].
Mary 3F.SG.IPFV-want 3F.SG.IPFV-go.out at.once
'Mary wants to go out at once.'

c. **Pawlu j-ista'* [***li*/*jekk*/*biex*** *j-oqgħod għal-l-eżami*].
Paul 3M.SG.IPFV-can **that/if/to** 3M.SG.IPFV-stay for-DEF-exam
'Paul can that/if/to he sits/sit for the exam.'

d. **Marija t-rid* [***li*/*jekk*/*biex*** *t-oħroġ minnufih*].
Mary 3F.SG.IPFV-want **that/if/to** 3F.SG.IPFV-go.out at.once
*'Mary wants that/if/to she goes/go out immediately.'

In this case, the implied subject of the embedded predicate must be coreferential with the matrix subject. Moreover, the embedded verb must be in the imperfect.

(87) a. **Pawlu j-ista'* [*n-oqogħd-u għal-l-eżami*].
Paul 3M.SG.IPFV-can 1.IPFV-stay-PL for-DEF-exam
*'Paul can we sit for the exam.'

b. **Pawlu j-ista'* [*qagħad għal-l-eżami*].
Paul 3M.SG.IPFV-can 3M.SG.PFV.stay for-DEF-exam
*'Paul can he sat for the exam.'

Note that *ried* can have either modal ('have to') or volitional ('wish to') meaning. When *ried* has volitional meaning, it must obligatorily occur with *li*.

6.6.2 Phasals

The following are three examples of phasal predicates: *beda* 'begin', *waqaf* 'stop', *kompla* 'continue'.

(88) *Tumas beda* [*j-irranġa d-dar*].
Thomas 3M.SG.PFV.begin 3M.SG.IPFV-arrange DEF-house
'Thomas started redoing the house.'

The same restrictions in terms of coreferentiality and tense/aspect as for the modals apply here.

(89) a. **Tumas beda* [*rranġa d-dar*].
Thomas 3M.SG.PFV.begin 3M.SG.PFV.arrange DEF-house
*'Thomas began he fixed the house.'

b. **Tumas beda* [*n-irranġa-w id-dar*].
Thomas 3M.SG.PFV.begin 1.IPFV-arrange-PL DEF-house
*'Thomas began we fix the house.'

6.6.3 Pseudo-predicates

The class of *pseudo-verbs/predicates* (see Fabri 1993, Peterson 2007 for details) includes *għoddu* 'be almost to', *għadu* 'still be', *qisu* 'seem to', and *donnu* 'seem to'. Their distinguishing feature is that they agree with the sentence subject with respect to person, number and gender through direct object clitics and they lack any other inflectional markings. Indeed, these constructions beg the question of how we define subject in these cases, but we will not pursue this point here (see Peterson 2007 for a discussion). Note that in the examples below, we avoid glossing such object clitics to circumvent this problem.

(90) a. *It-tifla* *għad-ha* [*t-aħsel* *il-kelb*].
DEF-girl still.be-3F.SG 3F.SG.IPFV-wash DEF-dog
'The girl is still washing the dog.'
b. *It-tifla* *għodd-ha* [*ħasl-et* *il-kelb*].
DEF-girl be.almost.to-3F.SG wash-3F.SG-PFV DEF-dog
'The girl has almost washed the dog.'
c. *It-tifla* *qis-ha* [*ħasl-et* *il-kelb*].
DEF-girl seem.to-3F.SG wash-3F.SG.PFV DEF-dog
'The girl seems to have washed the dog.'

The following example shows the typical use of the clitic as a pronominal direct object.

(91) *It-tifl-a* *ra-t-hom.*
DEF-girl- F.SG see-3F.SG.PFV.SBJ-3PL.OBJ
'The girl saw them.'

6.6.4 Impersonal verbs

Impersonal verbs occur only in the third person masculine singular, which is arguably the unmarked form. Just like pseudo-predicates, subject agreement takes place by means of obligatory object (either direct or indirect) clitics. Examples are *irnexxielu* 'manage to', *fadallu* 'still have to', *fettillu* 'happen to' and modal *messu* 'should'.

(92) *It-tifla* *rnexxie-lha* *t-aħsel* *il-kelb.*
DEF-girl manage.3M.SG.PFV-3F.SG 3F.SG.IPFV-wash DEF-dog
'The girl managed to wash the dog.'

The following example shows the typical use of the clitic as a pronominal indirect object.

(93) *It-tifel* *bagħt-ilha* *l-ktieb.*
DEF-boy send.3M.SG.PFV-3F.SG.IO DEF-book
'The boy sent her the book.'

The following are other examples.

(94) a. *It-tifla* *fadal-lha* [*t-aħsel* *il-kelb*].
DEF-girl still.have.to-3F.SG 3F.SG.IPFV-wash DEF-dog
'The girl still has to wash the dog.'
b. *It-tifla* *fettil-lha* [*t-aħsel* *il-kelb*].
DEF-girl happen.to-3F.SG 3F.SG.IPFV-wash DEF-dog
'The girl happened to wash the dog.'
c. *It-tifla* *qabil-lha* [*t-aħsel* *il-kelb*].
DEF-girl be.better.for-3F.SG 3F.SG.IPFV-wash DEF-dog
'It was better for the girl to wash the dog.'

6.6.5 Equi and raising predicates

A fifth type of predicate that occurs without a complementizer is the control predicate type, namely, equi and raising predicates (see also Section 4.2.4).

An example of an equi verb is *ipperswada* 'persuade'.

(95) a. *Il-ġuvni* *pperswada* *li-t-tifla* [*t-aħsel* *il-kelb*].
DEF-young.man persuade-3M.SG.PFV CSE-DEF-girl 3F.SG.IPFV-wash DEF-dog
'The young man persuaded the girl to wash the dog.'
b. **Il-ġuvni* *pperswada* *li-t-tifla* [***li*** *t-aħsel* *il-kelb*].
DEF-young.man persuade-3M.SG.PFV CSE-DEF-girl **that** 3F.SG.IPFV-wash DEF-dog
*'The young man persuaded the girl that she washes the dog.'

Other predicates of this type are *sforza* 'force' and *għallem* 'teach'.

Note that, in the case of *ipperswada* and *għallem*, *li* is not possible when the complement predicate encodes an activity but it is possible with the expression of non-activities, such as a state or a modality.

(96) a. *Il-ġuvni* *pperswada* *li-t-tifla* [***li*** *għand-ha*
DEF-young.man persuade.3M.SG.PFV CSE-DEF-girl **that** has-3F.SG
talent *sabiħ*].
talent nice
'The young man convinced the girl that she has good talent.'
b. *Il-ġuvni* *pperswada* *li-t-tifla* [***li*** *t-ista'*
DEF-young.man persuade.3M.SG.PFV CSE-DEF-girl **that** 3F.SG.IPFV-can
t-ieħu *sehem*].
3F.SG.IPFV-take part
'The young man convinced the girl that she could take part.'

Also, with *sforza* and *ipperswada*, although *li* is not possible with an activity, *biex* can occur as a complementizer (see Section 4.2.2).

(97) *Il-mara* *sfurza-t* *li-t-tifla* [***biex*** *t-aħsel* *il-kelb*].
DEF-woman force-3F.SG.PFV CSE-DEF-girl **to** 3F.SG.IPFV-wash DEF-dog
'The woman forced the girl to wash the dog.'

Finally, another example of the equi type is intransitive *ried* 'want', which does not take *biex*.

(98) a. *Il-mara* *t-rid* [*t-aħsel* *il-kelb*].
DEF-woman 3F.SG.IPFV-want 3F.SG.IPFV-wash DEF-dog
'The woman wants to wash the dog.'
b. **Il-mara* [*t-rid* ***biex*** *t-aħsel* *il-kelb*].
DEF-woman 3F.SG.IPFV-want **to** 3F.SG.IPFV-wash DEF-dog

Raising (to object) predicates include transitive *ried* 'want', *xtaq* 'wish', *stenna* 'expect', and immediate perception verbs like *ra* 'see' and *sema*' 'hear'. (99a) is the raised version of *ried* and (99b) is the *in situ* (unraised) version.

(99) a. *Il-mara* *ried-et* [*lil-l-ġuvni* *j-aħsel* *il-kelb*].
DEF-woman want-3F.SG.PFV CSE-DEF-young.man 3M.SG.IPFV.wash DEF-dog
'The woman wanted the young man to wash the dog.'
b. *Il-mara* *ried-et* [***li*** *l-ġuvni* *j-aħsel* *il-kelb*].
DEF-woman want-3F.SG.PFV **that** DEF-young.man 3M.SG.IPFV.wash DEF-dog
'The woman wanted that the young man wash the dog.'

Note that, in (99a) and (99b) desiderative *ried* requires a non-factive complement and the embedded verb has to be in the unmarked, imperfect form.

Things are slightly different with another predicate that can occur in a raised and non-raised version, namely, immediate perception *ra* 'see'. In (100a) and (100b) the complement verb has to express a situation either contemporaneous with, or prior to, the time of activity encoded by the matrix verb.

(100) a. *Il-mara* *ra-t* [*lil-l-ġuvni* *j-aħsel* *il-kelb*].
DEF-woman see-3F.SG.PFV CSE-DEF-young.man 3M.SG.IPFV-wash DEF-dog
'The woman saw the young man wash the dog.'
b. *Il-mara* *ra-t* [***li*** *l-ġuvni* *kien* *qed*
DEF-woman see-3F.SG.PFV **that** DEF-young.man be.3MS.G.PFV PROG
j-aħsel *il-kelb*].
3M.SG.IPFV-wash DEF-dog
'The woman saw that the young man was washing the dog.'

With respect to the examples discussed in this section, i.e., predicates which obligatorily occur without an explicit complementizer, there must be corefentiality between the understood subject of the complement and either the matrix

object (for transitive predicates) or subject (for intransitive predicates). In contrast, the occurrence of *li* does away with this referential dependency.

7 Conclusion

We have identified three main subordinating conjunctions in Maltese which can function as complementizers, namely, *li* (Section 4.1), *jekk* (Section 4.2.1) and *biex* (Section 4.2.2). These are in complementary distribution with a number of other expressions, typically question words such as *xi* 'what', *meta* 'when', *x'ħin* 'what time', *kif* 'how', *kemm* 'how much/many' (Section 4.2.3), which can also introduce a complement, and can, therefore, be assumed to occupy the complementizer position syntactically. We have also looked at cases in which the complement clause lacks an explicit complementizer (Section 4.2.4). We have also discussed the distinction between clauses functioning as complements or as adverbial modifiers (Section 5) and described in detail the distribution of the complementizer with respect to the different types of complement-taking predicates (Section 6).

A number of factors play a role in determining which complementizer occurs in a particular construction, namely, the semantics of the complement-taking predicate, coreferentiality and, within the complement clause, tense/aspect, state/action and factive/non-factive distinctions. *jekk*, which as a complementizer always requires the embedded verb to display the suffix *-x*, introduces a non-factive complement that characteristically expresses an element of uncertainty (Section 4.2.1). It can always be omitted as long as the suffix *-x* appears on the complement verb. *biex* always introduces a purpose clause, both as complementizer and as a subordinating conjunction; however, in the former but not in the latter case, it is optional (sections 4.2.2 and 6.2). *li*, on the other hand, cannot be as straightforwardly characterized as *jekk* and *biex*, since, in general, it occurs in contexts in which *jekk* and *biex* do not figure. It occurs easily in both factive and non-factive contexts and does not carry a specific meaning (e.g., uncertainty, purpose). In this sense, we conclude that *li* can be considered the default complementizer, as opposed to the more specific *jekk* and *biex*.

This is a first attempt at a detailed characterization of the semantics of complementizers in Maltese. We have provided a general picture of the facts, which have allowed us to give a first approximation of the functions and meanings of the main complementizers, and related issues. Further studies are required, in particular, to explore the difference between complement and modifying clauses in Maltese within a more general theoretical framework.

References

Aquilina, Joseph. 1959. *The Struture of Maltese*. Malta: The Royal University of Malta.

Borg, Albert. 1994a. Complementation in Maltese. In Joseph Brincat (ed.), *Languages of the Mediterranean,* 218–227. Malta: The Institute of Linguistics, The University of Malta.

Borg, Albert. 1994b. Maltese complement sentences to non-verbs. In Dominique Caubet & Martine (eds.), *Actes des premières journées internationales de dialectologie arabe de Paris*. 111–120. Paris: INALCO.

Borg, Albert & Marie Azzopardi-Alexander. 1997. *Maltese. Routledge Descriptive Grammars.* London: Routledge.

Borg, Albert & Marie Azzopardi-Alexander. 2009. Topicalization in Maltese. In Bernard Comrie, Ray Fabri, Elisabeth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese Linguistics,* 71–81. Amsterdam/Philadelphia: John Benjamins.

Borg, Albert & Manwel Mifsud. 2002. Maltese object marking in a Mediterranean context. In Paolo Ramat & Thomas Stolz (eds.), *Mediterranean Languages*, 33–46. Bochum: Brockmeyer Verlag.

Borg, Alexander. 1997. Maltese Phonology. In Alan S. Kaye (ed.), *Phonologies of Asia and Africa (Including the Caucasus)*, 245–288. Winona Lake, Indiana: Eisenbrauns.

Borg, Donna. 2011. *Il-Modalità Epistemika fil-Malti: Analiżi ta' xi Verbi u Espressjonijiet Verbali* [Epistemic modality in Maltese: an analysis of some verbs and verbal expressions]. Malta: University of Malta M.A. Thesis.

Brincat, Joseph. 2011. *Maltese and Other Languages: A Linguistic History of Malta*. Malta: Midsea Books.

Camilleri, Maris. 2009. *Clitics in Maltese*. Malta: University of Malta, B.A. (Hons) Thesis.

Camilleri, Maris. 2011. Relative constructions in Maltese. In Chris Cummins, Chi-Hé Elder, Thomas Godard, Morgen Macleod, Elaine Schmidt & George Walkden (eds.). *Proceedings of the Sixth Cambridge Postgraduate Conference in Language Research,* 29–42. Cambridge: Cambridge Institute of Language Research.

Comrie, Bernard. 1982. Syntactic-morphological discrepancies in Maltese sentence structure. *Communication & Cognition* 15(3/4). 281–306.

Dessoulavy, Charles L. 1938. *A Maltese-Arabic Word-List*. London: Luzac & Co.

Fabri , Ray. 1993. *Kongruenz und die Grammatik des Maltesischen.* Tübingen: Max Niemeyer Verlag.

Fabri, Ray. 1996. The Construct State and the Pseudo-Construct State in Maltese. *Rivista di Linguistica*, 8(1). 229–244.

Fabri, Ray. 2009. Stem allomorphy in the Maltese verb. In: *Ilsienna – Our Language* 1. 1–20.

Fabri, Ray. 2010. Maltese. In: Christian Delcourt and Piet van Sterkenburg (eds.) *The Languages of the New EU Member States* (Revue Belge de Philologie et d'Histoire 88[3]), 791–816.

Fabri, Ray & Albert Borg. 2002. Topic, Focus and Word Order in Maltese. In Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi & Zakia Iraqui-Sinaceur (eds.) *Aspects of the Dialects of Arabic Today,* 354–363. Rabat: Amapatril.

Kiss, Katalin É. (ed.). 1995. *Discourse Configurational Languages*. Oxford: Oxford University Press.

Maas, Utz. 2007. Complex predicates in Maltese: From a Neo-Arabic perspective. In Bernard Comrie, Ray Fabri, Elisabeth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove

(eds.), *Introducing Maltese Linguistics,* 113–132. Amsterdam & Philadelphia: John Benjamins.

Mifsud, Manwel. 1995. *Loan Verbs in Maltese: A descriptive and Comparative Study.* Leiden: E.J. Brill.

Noonan, Michael. 2007 [1984]. Complementation. In Timothy Shopen (ed.) *Language Typology and Syntactic Description*, vol. 2*, Complex Constructions,* 2nd edn.*,* 52–150. Cambridge: Cambridge University Press.

Peterson, John. 2007. "Pseudo-verbs": an analysis of non-verbal (co-)prediction in Maltese. In Bernard Comrie, Ray Fabri, Elisabeth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese Linguistics,* 181–204. Amsterdam & Philadelphia: John Benjamins.

Spagnol, Michael. 2011. *A tale of two morphologies: verb struture and argument alternations in Maltese.* Konstanz: University of Konstanz PhD Thesis.

Stolz, Thomas. 2009. Splitting the verb chain in modern literary Maltese. In Bernard Comrie, Ray Fabri, Elisabeth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese Linguistics,* 133–179. Amsterdam & Philadelphia: John Benjamins.

Petar Kehayov

Complementation marker semantics in Finnic (Estonian, Finnish, Karelian)

1 Introduction

According to the traditional classification (e.g. Laanest 1982), the Finnic branch of Finno-Ugric language family comprises seven languages: Finnish, Karelian, Veps, Ingrian, Votic, Estonian and Liv, three of which – Votic, Liv and Ingrian – are moribund or highly endangered. Of the remaining four languages – Estonian, Finnish, Karelian and Veps, the first three were systematically examined with respect to their complementation marker systems. In particular, this study is based on Standard Estonian (henceforth referred to as *Estonian*), Standard Finnish (henceforth referred to as *Finnish*), the Northern Karelian dialect of Karelian, the Olonets dialect of Karelian, and Lude, which is a transitional dialect between Karelian and Veps (Turunen 1946: 1–9), and whose complementation system can be considered to be quite similar to that of Veps, the fourth surviving Finnic language, not considered in this study.

Finnish has slightly over five million speakers, Estonian about one million, Northern Karelian between fifteen and twenty thousand, Olonets between ten and fifteen thousand, and Lude less than thousand speakers.[1] While Estonian and Finnish have undergone intensive and long-lasting contact with the Germanic languages (Low and High German, and Swedish, respectively), the dominant language in the area in which Karelian has traditionally been spoken is Russian.

The analysis of the Estonian and Finnish complementation marker systems is based on previous research on the topic, grammar and dictionary information, as well as on my own linguistic expertise and the native intuition of my Estonian and Finnish consultants. The data from the Karelian varieties was mostly collected during fieldwork (July 2011) in the Northern Karelian village of Jyškyjärvi (Russ *Yushkozero*) and in the surroundings of the town of Priäžä (Russ *Pryazha*), where I had several native speakers of Northern Karelian, Olonets and Lude at my disposal. The elicitation procedure involved sentences which were given to the consultants for a translation into their native variety. In addition to this, I consulted all relevant dictionaries and grammars of these Karelian varieties and collected Northern Karelian and Olonets material from the newspapers *Vienan*

1 These are very rough estimations, as the speech communities of Karelian dialects are rapidly diminishing.

Karjala and *Oma Mua*, which are written in the standard versions of Northern Karelian and Olonets.

Finnic languages posses a remarkable number of complementation marking devices; consider for instance the highlighted morphemes in examples (1)–(3) from Finnish.

(1) *Poliisi* *tietää,* ***että*** *poika* *on* *alaikäinen.*
police.NOM know:PRS.3SG **that** boy.NOM be.PRS.3SG underage.NOM
'The police know that the boy is underage.'

(2) *Poliisi* *tietää,* *on=**ko*** *poika* *alaikäinen.*
police.NOM know:PRS.3SG be.PRS.3SG=**PQ** boy.NOM underage.NOM
'The police know if the boy is underage.'

(3) *Poliisi* *tietää* *pojan* *ole-**van*** *alaikäinen.*
police.NOM know:PRS.3SG boy:GEN be-**PRS.PTCP:GEN** underage.NOM
'The police know of the boy being underage.'

The Finnish word *että* in (1) is the functional equivalent of the English complementizer *that*, and the polar question clitic *=ko* (or *=kö*, depending on the vowel harmony rule) attached to the verb in (2) is the functional equivalent of the English complementizer *if*. The complement predication in these two sentences is expressed by a finite verb, and therefore the markers *että* and *-ko*/*-kö* could be said to function as finite clause complementizers. In contrast, the complement predication in (3) contains a non-finite verb form, namely a participle occurring in a special construction, where the actor argument takes the Genitive case and not the nominative as in (1) and (2).

Conforming to the terminology agreed upon in the present volume, all highlighted items in the examples above will henceforth be called *complementation markers*. The subset of complementation markers which co-occur with a finite complement clause (as in [1] and [2]), will henceforth be referred to as (*canonical*) *complementizers*, whereas participle, infinitive, and nominalization markers requiring differential marking of their first argument and used for the coding of complements (as in [3]) will be called *non-finite complementation markers*.

A reader familiar with Finnish or Estonian grammar would note that this nomenclature is at odds with the "official" terminology. Descriptive grammars often do without terms like "complementizer" or "complementation marker". For instance, the most comprehensive grammar of Finnish calls *että* a "general conjunction" (VISK: §1144), and the academic grammar of Estonian calls its Estonian equivalent *et* a "subordinating conjunction" (EKG II: 110). Moreover, recent studies based on spoken Finnish and Estonian, especially those carried out within the framework of interactional linguistics, challenge both the "sub-

ordinating" and the "conjunction" status of these items (e.g. Laury & Seppänen 2008; Keevallik 2010). The conceptual confusion seems to be due to the different research questions: while descriptive grammarians and interactional linguists have been attempting to identify the invariant meaning of certain linguistic elements, the aim of the present study is to find out the semantic import of these elements when they occur in a complementation context.

However, the problem concerns not only the markers themselves, but also the status of the dependent syntactic unit which they arguably link to the main clause. We could in principle delimit the extension of the terms "complementizer" or "complementation marker" by considering only heads of syntactic units that can be described as "complement clauses", and disregard heads of complements belonging to lower layers of clause structure (such as constituent complements). In this case, however, one has to nail down the denotatum of a "complement clause". Thompson (2002), for instance, completely denies the existence of a "complement clause" as a distinct type of a clause.[2] Furthermore, even categories such as "clause" or "sentence" are better accounted for as gradual and not as discrete categories: different structures may manifest different amount of prototypical clausal or sentential features (see Lehmann 1988; VISK: §872).

In view of the lack of consensus concerning the basic notions necessary for a study on complementation, I will assume as wide as possible a definition of the notion "complementation marker", thus leaving space for further discrimination in the field. This definition is a somewhat expanded version of Noonan's (2007: 54) definition of complementizers. The label 'complementation marker' will stand for all words, particles, clitics, affixes and constructions combining words, particles, clitics, affixes and word order configurations, which function as syntactic templates used to identify an entity as a complement.

2 Inventory

Applying the terminology outlined above, we can distinguish between two types of complementation markers in Finnic: a) canonical finite clause complementizers (mostly subordinating conjunctions), and b) non-finite constructions serving as complementation markers. In the first case, recall examples (1) and (2), it is

2 Thompson (2002: 127) states: "the data strongly confirm that the syntactic behavior of the commonly recognized complement types for English (e.g., infinitives, *that*-complements, gerunds, and question-words) is highly disparate, and, to my knowledge, there exists no evidence to support their being considered to be the same category."

essentially a subordinator that serves to identify a syntactic unit as a complement. In the second case, recall example (3), it is a non-finite form of the dependent verb together with its argument coding properties that serves to identify a syntactic unit as a complement.

This section will present all complementation markers belonging to these two classes in Estonian, Finnish and Karelian. The survey of each type of complementation markers will concentrate more on formal than on functional observations. The latter will be discussed in Section 5, which addresses the semantic functions of complementation markers.

2.1 Canonical complementizers

Here one can further distinguish between two types of complementizers. The first includes subordinators which alone serve to identify a finite clause as a complement. In such cases it is only the subordinator, and no other property of the sentence, which identifies the clause as a complement. This type could be labelled *free complementizer*. The second type includes subordinators which identify finite clause as a complement, but only in combination with the Conditional mood[3] on the dependent predicate. It could be argued that in such cases the subordinator, which functions as a head of the finite complement clause triggers (or attracts) Conditional mood marking on the complement verb. This type could be labelled *complementizer with mood concord*. Despite the fact that it is not only the overt subordinator that identifies the complement clause as such here, the contribution of the subordinator is crucial: it occurs in the same position as free complementizers do (in the beginning of the dependent clause), and, moreover, there are chances that it develops into a free complementizer after some relaxation of the concord rule (a case in point will be discussed in Section 2.1.2.1).

2.1.1 Free complementizers

2.1.1.1 General complementizers (or THAT-complementizers)

The Estonian subordinator *et*, the Finnish *että* and Northern Karelian *jotta*, all etymologically related (Karelson 1959: 100–110), function as rough equivalents

3 The so-called "Conditional mood" in Finnic roughly corresponds to the Subjunctive in the Romance and in some Germanic languages (see e.g. Metslang 1999 for Finnic and Nordström 2010: 131–147 for Germanic), and as such corresponds to the typological mood irrealis (cf. Palmer 2001: 148–150).

of the English complementizer *that*. They could be called general complementizers as they are semantically neutral, i.e. they do not have impact on the meaning of the complement proposition. The truth value of the complement propositions of THAT-complementizers depends on the semantics of the matrix verb which means that they are semantically unmarked (see e.g. EKG II: 283 for Estonian, and Boye, van Lier & Brink 2015 for a cross-linguistic overview of such unmarked complementizers).

Other reasons for considering them as general complementizers are their frequency and the formal unmarkedness of their complements. They occur in a larger choice of syntagms than any other complementation marker and they take a complement clause which is identical with a declarative independent clause; compare (4) and (5) from Estonian.

(4) *Avastasin,* [***et*** *ta* *on* *juba* *kodus.*
discover:PST:1SG **that** s/he be.PRS.3SG already home:INE
'I discovered that s/he is already at home.'

(5) *Ta* *on* *juba* *kodus.*
s/he be.PRS.3SG already home:INE
'S/he is already at home.'

These etymologically related words function as THAT-complementizers in the majority of Finnic languages; in the eastern part of the Finnic area they have been replaced, however, by other subordinators. In Olonets and marginally in Lude, the temporal/conditional word *ku* 'when; if' is used in this function; see (6) from Standard Olonets.

(6) *En* *smiettinyh,* [***ku*** *kniigu* *roih* *muga*
NEG:1SG anticipate:PST.ACT.PTCP **that** book:PART get:PRS.3SG so
korgiesti *arvostettu*].
highly appreciate:PST.IMPS.PTCP
'I could not anticipate that the book would receive such high acclaim.'
(Oma Mua no. 26(1064), 06.07.2011, page 3)

In addition, Olonets and Lude tend to use the THAT-complementizer *što*, which is borrowed from Russian; cf. (7) from Lude.

(7) *Mina* *dūmaičen,* [***što*** *Īvan* *tuлou*].
I think:PRS.1SG **that** Ivan come:PRS.3SG
'I think that Ivan will come.'
(MT[4])

2.1.1.2 Q-complementizers

The Finnic languages have direct question markers occurring in complementizer position. In Estonian for instance, the polar question word *kas* 'whether, if' and the wh-question words function as complementizers; see (8) and (9).

(8) *Jaan* *küsis,* [***kas*** *Mari* *tuleb*].
Jaan ask:PST.3SG **PQ** Mari come:PRS.3SG
'Jaan asked if Mari is coming.'

(9) *Mul* *ei* *ole* *aimugi,* [***kuhu*** *lähen*].
I:ADE NEG be.CN clue:NEG.CL **where.ILL** go:PRS.1SG
'I don't have a clue where I am going.'

Finnish and Karelian have a special clitic for polar questions (=*ko*/=*kö*), which is attached to the first constituent of the complement clause. Usually this is the complement verb, the core of the proposition (recall example [2]), but in principle any other constituent could be addressed; cf. (10) from Finnish. For the expression of wh-complements, Finnish and Karelian, just like Estonian, use question words; see (11), also from Finnish.

(10) *Janne* *kysyi,* [*tänään=**kö*** *tulee* *Mari*].
Jaan ask:PST.3SG today=**PQ** come:PRS.3SG Mari
'Jaan asked if Mari is coming TODAY.'

(11) *Minulla* *ei* *ole* *aavistustakaan,* [***mihin*** *olen* *menossa*].
I:ADE NEG be.CN clue:NEG.CL **where:ILL** be:PRS.1SG going:INE
'I don't have a clue where I am going.'

In all three languages such complements are structurally identical with direct questions.[5]

4 Here and elsewhere I refer to the personal names of the consultants using abbrevaitions consisting of the initial letters of their first and last name.

5 This study is primary concerned with the polar type of Q-complementizers; the wh-type is only of marginal interest, as wh-words occur as arguments of the complement verb and therefore their complementizer status is debatable.

2.1.1.3 Temporal and conditional words as complementizers (or TEMP/COND-complementizers)

English encodes temporal clauses (cf. *I'll tell him* [*when he comes*].) and conditional protases (cf. *I'll tell him* [*if he comes*].) with different conjunctions. In contrast, Estonian and Olonets use a single conjunction both in temporal and conditional clauses. Furthermore, this conjunction (*kui* in Estonian and *ku* in Olonets[6]) is employed also as a complementizer. Consider example (12) from Estonian where the clause headed by *kui*, despite its possible temporal or conditional interpretation, fills the only argument slot of the evaluative predicate 'it is sad'.

(12) *On* *kurb,* [***kui*** *inimene* *oma* *juured* *kaotab*].
be.PRS.3SG sad **if/when** person his/her roots lose:PRS.3SG
'It is sad when/if a person loses his/her roots.'

On the other hand, Finnish, Northern Karelian and Lude, just like English, use different conjunctions as temporal and conditional adverbializers: in Finnish *kun* and *jos*, respectively. Of these two, the temporal adverbializer *kun* is more often employed as complementizer; cf. (13) from Finnish.

(13) *Riittää,* [***kun*** *sinä* *tulet*].
suffice:PRS.3SG **when** you come:PRS.2SG
'It's enough that you come' (lit. It suffices when you come).
(Sands 2000: 130)

It could be argued that the conjunction *kun* in (13) is ambiguous with respect to temporal and complementizer interpretations (cf. Herlin 1998: 75). Sands (2000: 131–132) provides, however, two pieces of evidence showing clearly that complement clauses with *kun* behave differently from temporal clauses with *kun*. First, unlike temporal clauses, complement clauses with *kun* are fixed in sentence final position and cannot be moved; this is a typical feature of THAT-complements in Finnish, which does not occur with adverbial clauses (Korhonen 1993: 20–21). Second, unlike temporal clauses, complement clauses with *kun* cannot be substituted for a certain non-finite form of the verb, the Inessive form of the so-called second infinitive.

In some restricted contexts the Finnish conditional adverbializer *jos* can also be employed as complementizer (in some Finnish dialects this is more common; see Karelson 1959: 132 for examples); consider (14).

6 Olonets sometimes uses the Russian protasis marker *jesli* in conditional clauses.

(14) *Minä koetan* [***jos*** *jaksan nostaa sen*].
I try:PRS.1SG **if** be.able:PRS.1SG lift.INF it:GEN
'I'll try if/whether I am able to lift it.'
(Korhonen 1993: 128)

In accordance with the formal difference between temporal and conditional adverbializers in Finnish, Northern Karelian and Lude, I will further distinguish in these languages between TEMP-complementizers and COND-complementizers.

2.1.1.4 Other lexemes as free complementizers

Other words may occasionally appear as complementizers, although such occurrences are attested only in very specific contexts. Consider example (15) in which the Estonian adverb *kuni* 'until' introduces a subject complement clause of the verb 'last, take time'.

(15) *Kestis kaua,* [***kuni*** *Peeter tuli*].
last:PST.3SG long **until** Peeter come:PST.3SG
'It took long time until Peeter came.' (lit. Until-Peeter-came took long time).
(Remmelg 2006: 17)

2.1.2 Complementizers with mood concord

All cases of mood concord in Finnic concern obligatory combinations of what is intuitively a complementizer (i.e. clause initial lexical subordinator) and a subjunctive-like mood (called "Conditional") on the complement verb.

2.1.2.1 Similative subordinators (or SIM-complementizers)

A finite clause is identified here as a complement by the combination of a lexeme with comparative-similative meaning, roughly corresponding to the English 'as if', and the Conditional mood morpheme on the dependent verb. In Estonian for instance, the nearly synonymous words *justkui*, *kui*, *justnagu*, *nagu*, and *otsekui*, all meaning 'as if; like; allegedly' require the Conditional morpheme *-ks* on the dependent verb; cf. (16).

(16) *On kuulda* [***justkui*** *ulu-ks hunt*].
be.PRS.3SG hear:INF **as.if** howl-COND.3SG wolf.NOM
'It sounds as if a wolf is howling.'

The corresponding Finnish expressions *aivan kuin, ihan kuin, ikään kuin*, and *kuin* function also as complementizers requiring the Conditional marker *-isi* on the dependent verb; cf. (17).

(17) *Välillä* *hän* *puhuu* [***ikään kuin*** *tämä* *koti* *ol-isi* *hänen*
sometimes s/he speak:PRS.3SG **as.if** this home be-COND.3SG his/her
vanhempiensa *koti*].
parents:PL.GEN:3POSS home
'Sometimes s/he talks as if this home were her/his parents' home.'

As already explained in the beginning of Section 2.1, in such cases the construction containing a lexical subordinator and non-indicative mood on the verb can be said to function as a complementizer. Another possible account for cases like (16) and (17) would be to say that the Estonian *justkui, kui, justnagu, kui*, and *otsekui* and the Finnish *aivan kuin, ihan kuin, ikäänkuin* and *kuin* are irrealis complementizers (cf. Nordström 2010: 121), which obligatorily trigger the Conditional mood on the complement verb.

In Karelian dialects, on the other hand, the Conditional is not obligatory, which means that the lexeme 'as if' has become independent of the mood marking on the verb. In Northern Karelian, the similative word *niinku* is used either with the Conditional or with the Indicative on the complement verb. In Olonets and Lude the Russian loanwords *buitta* and *rouno* (or *rovno*) are used with the temporal-conditional word *ku*, and the compound meaning of these words is 'as if'. Mirroring the Russian pattern of using the Indicative in such contexts (see Letuchiy 2008: 222), they usually occur with the Indicative on the complement verb; cf. (18) from Olonets with *rounoku*.

(18) *Näin* *unis* [***rownoku*** *kaksitoštu* *tedrie* *puzuh*
see:PST:1SG dream:INE **as.if** twelve blackcock:PART cage:ILL
katoin].
catch:IND.PST.1SG
'I dreamt that (lit. as if) I had caught twelve blackcocks in the cage.'
(SKJ-LD: 313)

Similar developments have taken place in Baltic. Holvoet (this volume) discusses the development of Lithuanian *lyg* and Latvian *it kā* 'as if' from adverbializers expressing counterfactive comparison and used only with irrealis forms of the complement verb, to complementizers occurring after speech act predicates and predicates of propositional attitude and marking that the content of the clausal

complement is contrary to the facts. Much like in Eastern Finnic, in such contexts the irrealis marking on the complement verbs is no longer compulsory.[7]

The necessary conjecture is that there is a tendency both in Finnic and in Baltic for the similative lexeme to gain autonomy from the irrealis marking on the complement verb. What we observe here is that a certain type of complementizer with mood concord changes to free complementizer.

Notwithstanding the foregoing, SIM-complementizers should be characterized as the most "coercive" complementizer type in Finnic. Apart from the mood concord, the Estonian *justkui* 'as if; allegedly' requires also changing of the word order of the complement clause from SV to VS; compare (19), with the general complementizer *et* and the unmarked SV order in the dependent clause, and (20) with the SIM-complementizer and VS order in the dependent clause.

(19) *Ta* *rääkis,* [***et*** *sa* *oled* *ta* *üksinda* *jätnud*].
s/he speak:PST.3SG **that** you be:IND.2SG her/him alone leave:ACT.PST.PTCP
'S/he said that you had left her/him alone.'

(20) *Ta* *rääkis* [***justkui*** *ole-ksid* *sa* *ta* *üksinda*
s/he speak:PST.3SG **as.if** be-COND:2SG you her/him alone
jätnud]
leave:ACT.PST.PTCP
'S/he pretended that you had left her/him alone'
(lit. S/he was saying as if you have left her/him alone).

This degree of coercion may be taken to suggest that the similative lexemes are less grammaticalized as complementizers than the other finite clause complementizers.

2.1.2.2 Other lexeme as a complementizer with mood concord

An example of a lexeme, which marginally functions as a complementizer, and obligatorily co-occurs with the Conditional mood, comes from Finnish. Consider the adverbializer *vaikka* 'although' in (21), where it introduces an object clause of knowledge predicate.

(21) *Trond* *ei* *tiedä* [***vaikka*** *läht-isi-n=kin*].
Trond NEG know.CN **although** go-COND-1SG=CL
'Trond doesn't know if/that maybe I'll go.' (Korhonen 1993: 128)

7 Contemporary colloquial Estonian shows signs of relaxing the Conditional concord rule as well; this was suggested to me by Liina Linström and Helle Metslang and could be confirmed by a search in Internet.

The combination of *vaikka* and the Conditional mood on the verb implies a contradiction between the expectations of Trond and desire of the protagonist of the complement clause to remain agnostic (Sands 2000: 131).

Table 1 summarizes the canonical complementizer types presented in this section.

Table 1: Canonical complementizers: an outline of the major types

functional type	model context	reflexes	mood concord
THAT-COMPLEMENTIZER	*Mary discovered that John is gone.*	Estonian *et*, Finnish *että*, *jotta*, Northern Karelian *jotta*, Olonets *što*, *ku*, Lude *što*	NO
Q-COMPLEMENTIZER			
a) POLAR	*John asked Mary if he could come.*	Estonian *kas*, Finnish *-ko/-kö*, Northern Karelian *-ko/-kö,* Olonets *-ko/-kö*, Lude *-ko/-kö*, *li*	NO
b) WH-	*Mary asked where her passport is.*	wh-word	NO
TEMP/COND-COMPLEMENTIZER	*It is sad when/if one loses his integrity.*	Estonian *kui*, Finnish *kun*, *jos*, Northern Karelian *kui, hos*, Olonets *ku*, Lude *ku/konz, d'jesl'i*	NO
SIM-COMPLEMENTIZER	*John dreamt as if he was the king of Scotland.*	Estonian *justkui*, *kui*, *justnagu*, *nagu*, *otsekui*, Finnish *aivan kuin*, *ihan kuin*, *ikäänkuin*, *kuin*, Northern Karelian *niinku*, Olonets/Lude *buittaku, rounoku*	YES NO

2.2 Non-finite complementation markers

This section deals with structures characterized by the lack of dedicated marker identifying the complement clause as such. Here only the grammatical properties of the dependent construction identify it as a complement. As already noted in Section 1, in such constructions the first argument of the dependent verb is encoded differently from the subject of the matrix verb. The degree of interlacing between clauses here is higher than the degree of interlacing between finite clauses linked by lexical subordinators (Lehmann 1988). Particular questions of interlacing, such as shared participants and raising will not be addressed in detail in the following discussion, as they do not have direct bearing on the semantics of non-finite complementation markers. Because of the interlacing (i.e. constitu-

ent sharing with the matrix clause) non-finite complements will not be presented in square brackets.

The Finnic non-finites functioning as complementation markers can be classified in three major types: a) participles, b) infinitives, and c) nominalizations.

2.2.1 Participle as a complementation marker

This type of complementation marking is found in all Finnic languages. The original present participle in Finnic is characterized by the suffix *-v(A)*, whereas the past participle is manifested by the suffix *-nUt* (see Laanest 1982: 162). In Finnish, the present (22) and the past participle (23) occur in the Genitive case, whereas in Estonian the participle is in the Partitive case; see (24) and (25).

(22) *Pelkäsin* *Jannen* ***sano-van*** *sen.*
be.afraid:PST:1SG Janne:GEN **say-PRS.PTCP:GEN** that:GEN
'I was afraid that Janne would say this.'

(23) *Pelkäsin* *Jannen* ***sano-neen*** *sen.*
be.afraid:PST:1SG Janne:GEN **say-PST.PTCP:GEN** that:GEN
'I was afraid that Janne has said this.'

(24) *Kartsin* *Jaani* *seda* ***ütle-vat.***
be.afraid:PST:1SG Jaan:PART this:PART **say-PRS.PTCP:PART**
'I was afraid that Janne would say this.'

(25) *Kartsin* *Jaani* *seda* ***öel-nuvat.***
be.afraid:PST:1SG Jaan:PART this:PART **say-PST.PTCP:*vat* (≈PRS.PTCP:PART)**
'I was afraid that Janne would say this.'

The Estonian Partitive participle is more grammaticalized than the Finnish Genitive participle. It is no more conceived by speakers as a participle, but as a special non-finite form ending in *-vat*, which occurs mainly in complements. This is obvious in (25) with the form *-nuvat*, in which the already opaque element *-vat* (once present participle in the Partitive) is added to the past participle: the past tense marker *-nu* of this participle and the present tense marker *-v*, which are otherwise incompatible, co-occur, because the latter is no longer perceived as a tense marker.

Karelian dialects have mostly substituted the inherited Finnic present participle for an agentive nominalization ending in *-jA* (see e.g. Zaikov 1992: 45; Markianova 1993: 46). In complements tensed past they use the original *nUt*-participle, just like Estonian and Finnish.

2.2.2 Infinitive as a complementation marker

Two infinitives usually encode complements: the first of them is manifested by the suffix *-A/DA*, the second by the suffix *-MA*. In the Finnish grammatical tradition these are called "first infinitive" and "third infinitive" respectively, but for the sake of clarity here they will be labelled according to their affixal exponents: *A/DA*-infinitive and *MA*-infinitive (consult Saukkonen 1965: 5–58 for a comprehensive overview of the Finnic infinitives).

2.2.2.1 A/DA-infinitive as a complementation marker

The complement is identified as such here by the verb occurring in the form of the *A/DA*-infinitive; cf. (26) from Finnish.

(26) *Janne* *ehdotti* *Marille* ***lähte-ä*** *kävelylle.*
Janne suggest:PST.3SG Mari:ALL **go-*A/DA*** walk:ALL
'Janne suggested to Mari to go for a walk.'

2.2.2.2 MA-infinitive as a complementation marker

Here the complement is marked by the *MA*-infinitive form of the complement verb. In Finnic, this infinitive occurs as a complementation marker in three local case forms: Illative (expressing movement inward), Inessive (expressing fixed location within something), and Elative (expressing movement outward).[8]

The most frequent case form of the *MA*-infinitive is the Illative. As a complementation marker, this form is in complementary distribution with the *A/DA*-infinitive: some verbs take complements in the form of *A/DA*-infinitive; other verbs, like 'ask, beseech' in (27) from Finnish, take complements in the Illative form of the *MA*-infinitive.

(27) *Janne* *pyysi* *Maria* ***lähte-mään*** *kävelylle.*
Janne ask:PST.3SG Mari:PART **go-*MA*.ILL** walk:ALL
'Janne asked Mari to go for a walk.'

The Inessive case form of the *MA*-infinitive is semantically more specific than the Illative; it conveys that the subject of the complement verb is in process of doing something, roughly corresponding to the English Progressive; cf. (28) from Finnish.

8 The complement infinitive also occurs marginally in the Instructive case, which has instrumental function in the nominal inflection.

(28) *Kuuntelin* *naapuria* ***riitele-mässä*** *vaimonsa* *kanssa*.
listen:PST:1SG neighbour:PART **argue-*MA*:INE** wife:GEN.3POSS with
'I was listening to the neighbour arguing with his wife.'

The Elative case form of *MA*-infinitive usually conveys that the subject of the complement clause is moving away from the process of doing something. It occurs after contra-implicative verbs, in other words, after verbs implying that the activity described in the complement did not take place (Sands 2000: 141; EKG II: 257); cf. (29) from Estonian.

(29) *Jaan* *keelas* *Marit* *kodust* ***lahku-mast***.
Jaan forbid:PST.3SG Mari:PART home:ELA **leave-*MA*:ELA**
'Jaan forbade Mari to leave home.'

2.2.3 Nominalization as a complementation marker

In all Finnic languages the action nominal ending in *-min(en)* can mark complement clauses; cf. (30) from Finnish with the Illative case form of the action nominal.

(30) *Hän* *keskittyy* *parhaillaan* *muistelmiensa*
s/he concentrate:PRS.3SG at.the.moment memoir:PL.GEN:3POSS
kirjoitta-miseen.
write-ACNMLZ:ILL
'At the moment he is concentrating on writing his memoirs.'
(Sands 2000: 141)

This nominalization belongs historically to the infinitive system and still shares some properties with the infinitives.[9] Until recently it was called "the fourth infinitive" in Finnish grammars (cf. VISK: §492).

Table 2 presents a synopsis of the non-finite complementation markers discussed in this section.

9 For instance, it occurs after modal verbs; see examples in Saukkonen (1965: 36).

Table 2: Non-finite complementation markers: an outline of the major types

functional type	model context	reflexes
PARTICIPLE	*I was afraid of John saying this again.*	Estonian PRS *-vat*, PST *-nuvad,* Finnish PRS *-vAn*, PST *-neen*, Karelian PRS *-vAn*/*-vAt*, PST *-nUOt*/*-nUOn*
INFINITIVE		
a) *A/DA*-INFINITIVE	*John suggested to Mary to go for a walk.*	Estonian *-a/da/ta*, Finnish *-A/dA/tA*, Karelian *-A*/*O*
b) *MA*-INFINITIVE		
ILLATIVE	*John asked Mary to come with him.*	Estonian *-ma*, Finnish *-mA*, Karelian *-mA(h)*
INESSIVE	*I heard the neighbour arguing with his wife.*	Estonian *-mas*, Finnish *-mAssA*, Karelian *-mAššA*
ABLATIVE	*Mary forbade John to leave the house.*	Estonian *-mast*, Finnish *-mAstA*, Karelian *-mAštA*
ACTION NOMINAL	*I believe in his coming back.*	Estonian *-mine* (+ case endings), Finnish *-minen* (+ case endings), Karelian *-mine/-mini* (+ case endings)

3 On the status of the complement

It could be argued that the complements discussed in the last section do not contain finite verbs and therefore cannot be regarded as clauses. Accordingly, if we chose to define "complementation markers" as heads of syntactic units that can be described as "complement clauses", we should disregard all elements introducing non-clausal complements. In this section I will elaborate on the functional weight of the non-finite complements enlisted in Section 2.2, and particularly on the question which of them can be considered to be clauses.

Clauses typically express propositions (truth-valued meaning units), have independent tense, can be modified by epistemic adverbs and occur after verbs of propositional attitude (such as 'believe') (cf. Lyons 1977: 445; Dik & Hengeveld 1991; Dixon 2006: 15–16; see also Boye 2010a for discussion). A dependent clause describes a situation with participants coded as arguments of its predicate. The first argument of the dependant predicate may but need not be coreferential with the subject of the matrix predicate.

In Finnic, finite *that*-complements manifest all clausal features listed above: they are tensed, they occur after propositional attitude verbs, they can be epistemically qualified and they can have non-logophoric subjects. Participial complements, on the other hand, do not have independent tense, but they do have

taxis (relative tense); see (31) from Finnish. The fact that participles can be tensed (past vs. non-past) means that participial complements do not require that the activity described by them is simultaneous with the activity described in the main clause (consult Dik & Hengeveld 1991: 240 on non-simultaneity as a property of propositional complements). Furthermore, as (31) and (32) from Finnish show, they occur after verbs of propositional attitude, they can be modified by epistemic adverbs, and they may have subjects which are not coreferential with the main clause subject.

(31) *Mari uskoo Jannen **kertovan** / **kertoneen** totuuden.*
Mari believe:PRS.3SG Janne:GEN **tell:PRS.PTCP:GEN** **tell:PST.PTCP:GEN** truth:GEN
'Mari believes that Janne is telling her the truth. / Mary believes that Janne has told her the truth.'

(32) *Mari arveli Jannen ilmeisesti **tulevan**, jos hänelle*
Mari suppose:PST.3SG Janne:GEN probably **come:PRS.PTCP:GEN** if he:ALL
tarjoittaisiin rahaa tästä hyvästä.
offer:IMPS.COND money:PART this:ELA good:ELA
'Mari supposed that Janne would probably come if he would be offered money for that.'

Infinitival complements are clearly less clausal than participial complements; consider the following examples from Finnish with the Illative case form of the *MA*-infinitive. Infinitives cannot be tensed, cannot be epistemically qualified (example [33]) or occur as complements of propositional attitude verbs (example [34]). In contrast, they can have a logical subject which is referentially distinct from the main clause subject; cf. (35).

(33) **Mari pyysi Jannea ilmeisesti **tulemaan**.*
Mari ask:PST.3SG Janne:PART probably **come:*MA*:ILL**
'*Mari asked Janne probably to come.'

(34) **Mari uskoo Jannea **tulemaan**.*
Mari believe:PRS.3SG Janne:PART **come:*MA*:ILL**
'*Mari believes Janne to come.'

(35) *Mari pyysi Jannea **tulemaan**.*
Mari ask:PST.3SG Janne:PART **come:*MA*:ILL**
'Mari asked Janne to come.'

Sentence (33) is grammatical only with a wide scope reading of the adverb *ilmeisesti* 'probably, apparently'. In this case, the probability of Mari asking Janne is evaluated and not the probability of Janne's coming.

Finally, as can be seen from the Finnish examples (36)–(38), action nominals cannot be tensed or epistemically modified, but they can occur with verbs

of propositional attitude and they may have an actor argument, which is referentially distinct from the subject of the main clause.

(36) **Häneen* *mahdollisesti* ***luottaminen*** *olisi* *virhe*.
s/he:ILL possibly **trust:ACNMLZ** be:COND.3SG mistake
'[?]Possibly trusting him/her would be a mistake.'

(37) *Mari* *uskoo* *Jannen* ***tulemiseen***.
Mari believe:PRS.3SG Janne:GEN **come:ACNMLZ:ILL**
'Mari believes in Janne's coming.'

(38) *Siivojan* ***osallistuminen*** *talonkokouksessa* *on* *varsin* *suositeltava*.
cleaner:GEN **participate:ACNMLZ** house.meeting:INE be.PRS.3SG quite recommendable
'The participation of the cleaning woman/man in the house meeting is strongly recommended.'

The attentive reader will have noticed that in (37), the Finnish verb *uskoa* 'believe' takes an action nominal complement in lative case (the so-called Illative in Finnish), its meaning being 'believe in something'. One may argue that here 'Janne's coming' is not in an object case, but in an adverbial case, and therefore this construction cannot be granted the status of a complement. Note however, that the lative argument should be analysed here not as oblique, but as extension-to-core argument (Dixon 2006: 7, 15). Besides, other verbs of propositional attitude (such as 'think') occur with action nominals encoded as canonical object arguments in the Partitive case; cf. (39).

(39) *Pojat* *miettivät* *tyttöjen* *naiseksi* ***tulemista***.
boys contemplate:PRS.3PL girl:PL:GEN woman:TRANS **come:ACNMLZ:PART**
'The boys reflect on girls becoming women.'

The restrictions governing the occurrence of the complement types presented above are summarized in Table 3.

Table 3: Compatibility of *that*-complements, participles, infinitives and action nominals with clausal criteria

	***that*-clause**	**participle**	**infinitive**	**action nominal**
tense	+	+/–	–	–
epistemic qualification	+	+	–	–
CTP of propositional attitude	+	+	–	+
non-logophoric subject	+	+	+	+

If we consider propositionality (i.e. truth-functionality) to be the main criterion for clausal status, it would probably be reasonable to regard *that*-clauses and participles as clausal complements and infinitives and nominalizations as non-clausal complements. However, the distribution in the table is not clear-cut and therefore I have chosen to approach the topic of this volume from across-the-board perspective.

4 Distribution of complementation markers

To prevent possible confusion I would like to clarify what I mean by distribution of complementation markers, on the one hand, and semantic functions of complementation markers, on the other. I will speak about distribution when the choice of certain complementation marker is determined by structural or semantic differences between sentences. I will speak about semantic functions when complementation markers, which occur with the same complement taking predicates (CTP-s) and in otherwise identical contexts can be analysed as having impact on the meaning of complements. Minimal pairs of sentences are the best heuristics for identification of semantic functions, as they allow extracting the semantic contribution of the complementation marker from the properties of the context.

The markers discussed in Section 2 could be studied with respect to various distributional criteria: for instance, they may be checked for the syntactic category of their complement (subject, object or other), for the constructional properties of the complement (e.g. for case marking on the arguments of the complement verb), for coreferentiality restrictions on arguments of the two clauses (see Lehmann 1988), for the position of the complement clause in relation to the main clause (preposed vs. postposed), or for the complexity of the complement (e.g. according to the number of constituents[10]). As none of these criteria seem to be of importance for the semantics of complementizers in Finnic, they will be disregarded.

The Finnic complementation markers were systematically studied only with respect to the complement taking predicates they occur with. The results are presented in Table 4, which takes three variables into account: a) the semantic type of the complement taking predicate, b) the type of complementation marker, and c) the number of verbs occurring or not occurring with the given type. A fourth variable could be the number of languages in which the given combination of a predicate and complementation marker type is attested. The languages studied

10 The non-finite types seem to be disallowed in complements with many constituents.

are, however, quite uniform in this respect; there are only few cases where languages use different types of complementation markers in the same context.[11] Having three (or four) variables clearly requires three- (or four-)dimensional representation, but as the distribution of complementizer types according to the CTP is only of secondary importance in this study, I will accept the limits of the two-dimensional display. Based on Noonan's classification (2007), I checked six cross-linguistically salient types of CTP-s for co-occurrence with complementation markers: two factive (or semi-factive[12]) and four non-factive.

A combination of a type of CTP and a complementation marker is given positive value in the table (+) if it occurs with the majority of the verbs listed under the given type of predicate. In case only half or the minority of the verbs listed under the given type occur with the given complementation marker, the combination is given the compromising value (–/+).

Table 4: Complementation markers according to the class of CPT (based on Noonan 2007)[13]

	Semantic type of the predicate	THAT-COMP	Q-COMP (polar)	TEMP/COND-COMP	SIM-COMP	*A/DA*-COMP	*MA*-COMP	PTCP. COMP	AC-NMLZ-COMP
(semi-)factive	immediate perception (*see, watch, hear, feel*)	+	+	+	–/+	–/+	+	+	+
	knowledge (*know*)	+	+	+	–	–	–	+	–
non-factive	utterance (*say, tell/report*)	+	+	–	+	–	–	+	–/+
	propositional attitude (*believe, suppose, guess*)	+	–	–	+	–	–	+	–/+
	desiderative (*want, hope*)	+	–	–	–/+	+	–	+	–/+
	manipulative (*command, request/ask*)	+	–	–	–	+	+	–	–

11 The only exception concerns differences in the distribution of the *A/DA*-infinitive and *MA*-infinitive.

12 Knowledge predicates and some perception predicates are not true factives, but semi-factives, in the sense of Karttunen (1971).

13 The distribution shown for the *MA*-infinitive and the action nominalization (NMLZ) is a sum of all case forms these structures exhibit in complements.

The distribution in the table allows for the following assumptions. The qualification *neutral* is justified for THAT-complementizers, as they occur with all predicate types. Another candidate for this qualification is the participial type, which is found in all predicate types except from the manipulative one. This is in line with the evidence from the history of Finno-Ugric languages; non-finite complements were predominant in the earlier stages of Finno-Ugric, and they still are in some eastern Finno-Ugric languages; see the discussion in Section 6. As participles are the main non-finite complementation device, they could be considered the major contestants of THAT-complementizers in the functional domain of complementation. Sands (2000: 127) noted that this is in stark contrast to what is known from the typological literature. Participial complements play only a rather restricted role in the complementation systems of languages worldwide, being usually limited to complements of immediate perception predicates (Noonan 2007). In Finnic, on the other hand, they have a wider use: in addition to the verbs in Table 4, participles also occur after suspicion verbs; cf. (40) from Finnish.

(40)	*Hän*	*epäilee*	*aviomiehensä*	***pettävän***	*häntä*.
	s/he	suspect:PRS.3SG	husband:GEN:3POSS	**cheat:PRS.PTCP:GEN**	she:PART

'She suspects that her husband has been cheating on her.'

The co-existence of two neutral complementation markers in one language suggests that they belong to two non-complementary sets of expressions, in other words, to two parallel complementation marking systems.

Now if we change the perspective and look at the distribution in Table 4 from the point of view of the predicate type, we may observe that the immediate perception verbs are almost indifferent to the type of complementation marker. While THAT-complementizers can be considered to be the neutral (or unmarked) complementation marker type with respect to the type of CTP, immediate perception predicates are the neutral type of CTP with respect to the choice of complementation marker.

Another minor observation concerning the distribution of the immediate perception verbs is that some of them occur with the SIM-complementizer, which is more or less reserved for non-factives. This shows simply that the similative type provides its own evaluation of the reality of the situation described in complement, which is independent of the factivity of the main verb. SIM-complementizers do not only mark the complement as irrealis (and as such tend to occur in harmonic combinations with irrealis forms of the complement verb, such as the Conditional), but also express negative propositional attitude toward the content of the complement. In (41) from Estonian, the speaker assumes that he has been deceived by his senses.

(41) *Ma tundsin* [***nagu*** *keegi roniks mööda mu selga*].
I feel:PST:1SG **as.if** somebody climb.up:COND.3SG along my back:PART.
'I felt as if somebody were climbing up my back.'

Lexical aspect can be also responsible for the distribution of complementation markers according to complement taking predicates. In Finnish, for example, the action nominalization and the *A/DA*-infinitive are by and large in complementary distribution (Sands 2000: 143; observe also the distribution in Table 4). In other words, these two complementation markers do not occur with the same verbs, and verbs allow only one of them. This is because the nominalization, according to Sands, "refers to a process over a period of time, the unfolding of activity, whereas the infinitive refers to the moment at which one gets involved in the activity". Thus, *A/DA*-infinitive is not compatible with verbs which describe a process, see (42a), "because the infinitive represents a moment in time" and, conversely, the nominalization is not compatible with verbs which relate to the beginning point of the activity, see (42b) (Sands 2000: 143).

(42) a. *Jatkan* [ok]*ruuan laittamista* / **laittaa ruokaa*.
continue:PRS.1SG food:GEN prepare:ACNMLZ:PART prepare:*A/DA* food:PART
'I continue preparing the food.'
d. *Aion *ruuan laittamista* / [ok]*laittaa ruokaa*.
intend:PRS.1SG food:GEN prepare:ACNMLZ:PART prepare:*A/DA* food:PART
'I intend to prepare the food.'

There are few verbs which allow for both complementation markers; such are the predicates of fearing, which seem to be underspecified for this aspectual distinction; cf. (43) and (44).

(43) *Janne pelkää nöyryytetyksi* ***tulemista***.
Janne be.afraid:PRS.3SG humiliated:TRANS **become:ACNMLZ:PART**
'Janne is afraid of being humiliated.'

(44) *Janne pelkää* ***tulla*** *nöyryytetyksi*.
Janne be.afraid:PRS.3SG **become.*A/DA*** humiliated:TRANS
'Janne is afraid of being humiliated.'

The cells on the horizontal axis in Table 4 which share positive values are candidates for semantic differentiation. They indicate that two or more complementation markers are used after the same complement taking predicate, and, thus, any possible semantic differences in minimal pairs of sentences with this predicate should be due to the switch of the complementation marker. However, different complementation markers may be found after the same predicates also for reasons other than semantic ones. For instance, the choice between THAT-

complementizers and the participial type is often related to the register, with a preference for the first in more informal registers and for the latter in higher, written varieties.

In the next section I will examine the distribution of complementation markers outside of complement clauses, as well as their omissibility and combinability. The discussion will be confined to the finite clause complementizers. Non-finite complementation markers cannot be omitted or combined, as they consist of obligatory morphological features without which the complement would not be grammatical, and which are specified only once per clause.

4.1 External distribution of complementation markers

Forms which are identified as complementizers may occur outside of what is considered to be the domain of complementation. There are two main issues that should be discussed in relation to such external distribution of complementizers. First, complementizers may occur in other types of dependent clauses, and, second, they may occur in independent clauses.

4.1.1 Complementizers in other types of dependent clauses

The free complementizers are generally employed in Finnic as subordinators in other types of dependent clauses.

THAT-complementizers, such as Estonian *et*, Finnish *että*, North-Karelian *jotta* and Olonets and Lude *što* and *ku* occur in various types of adverbial clauses (see e.g. EKG II: 295–311; Remmelg 2006: 14 for Estonian and VISK: §1111, §1160 for Finnish[14]). The most typical adverbializer uses of THAT-complementizers concern purpose and reason clauses; see the Olonets examples (45) and (46) with the complementizer *ku*.

(45) *Meil* *pidäü* *hüvin* *opastuo,* ***ku*** *parembah*
we:ADE must:PRS.3SG well study:*A/DA* **in.order.to** better
loppie *universitiettu*.
finish:*A/DA* university:PART
'Here one has to study well in order to finish the university with better success.'
(Markianova 1993: 93)

14 The standard variety of Finnish usually differentiates between *että* for complement clauses and *jotta* for adverbial clauses.

(46) *Hänel* *oli* *pahua* *mieldü,* ***ku*** *poigu* *ei*
s/he:ADE be:PST.3SG bad:PART mind:PART **because** boy NEG.3SG
tulluh *kodih.*
come:PST.PTCP home:ILL
'S/he was in a bad mood because the boy didn't come home.'
(Markianova 1993: 93)

Extensions of THAT-complementizers to relative clauses are, on the other hand, rare; cf. (47) from Estonian.

(47) *Ta* *on* *niisugune* *mees,* ***et*** *unustab* *kõik*
s/he.NOM be.PRS.3SG such.NOM man.NOM **that** forget:PRS.3SG everything
ära.
out
'He is the kind of a man who forgets everything.'
(Remmelg 2006: 14)

Q-complementizers are identical in form with direct question markers. Polar question complementizers do not occur in other dependent clauses except from complement ones. In Estonian and Karelian, but not in Finnish, wh-words function usually as relativizers and adverbializers; cf. (48) and (49) from Estonian

(48) *Ta* *otsib* *inimest,* ***kellest*** *saaks* *hoolitseda.*
s/he.NOM search:PRS.3SG person:PART **who:ELA** get:COND.3SG take.care:*A/DA*
'S/he is looking for a person whom s/he could take care of.'

(49) *Otsi* *sahtlist,* ***kus*** *on* *tema* *asjad!*
search.IMP.2SG drawer:ELA **where.NOM** be.PRS.3SG her/his stuff.NOM
'Look in the drawer where her/his stuff is!'

Finnish, on the other hand, uses in relative and adverbial clauses special series of conjunctions (with the stem *jo*-), which are formally distinct from wh-words.[15]

As already mentioned in Section 2.1.1.3, the TEMP/COND-complementizers occur mostly in temporal and/or conditional adverbial clauses. The similative type, which consists of a similative adverb ('as if; like') and the Conditional mood, is attested in comparative adverbial clauses.

(50) *Pääministeri* *käyttäytyy,* ***ikään kuin*** *vain* *hän*
prime.minister.NOM act:PRS.3SG **as.if** only he.NOM
tietäisi *asiasta* *parhaiten.*
know:COND.3SG thing:ELA best
'The prime minister acts as if he were the one who is best informed about the facts.'

15 However, in Colloquial Finnish wh-question words occur also as relativizers and adverbializers.

The only type restricted in subordination to complements is the polar Q-complementizer.

4.1.2 Complementizers in main clauses

Finnic complementizers occur in typical insubordination structures, where the formally subordinated clause is conventionalised as main clause (Evans 2007: 367); consider example (51) with the Northern Karelian general complementizer and (52) with the Estonian TEMP/COND-complementizer.

(51) ***Jotta*** *sie* *miun* *šilmistä* *katosit* *poikieš!*
that you mine eye:PL:ELA disappear:COND:2SG out
'Get out of my sight!'
(MK)

(52) ***Kui*** *ta* *ainult* *tuleks* *mulle* *külla!*
if/when s/he only come:COND.3SG I:ALL on.a.visit
'If only s/he would come to visit me!'

The (formally) indirect command in (51) and the if-clause in (52) represent typologically common types of insubordination structures (cf. Evans 2007).

On formal grounds, it could be argued that (51) and (52) are examples of elliptical main clauses containing manipulative and desiderative predicate.[16] This analysis is however impeded by the fact that both sentences clearly have an expressive interpretation, as signalled by the exclamation mark. They cannot be synchronically analyzed as complements of omitted main clauses for dependent clauses do not have illocutionary force (Lehmann 1988; Cristofaro 2003: 34) and thus cannot have expressive interpretation. The illocutionary force is signalled also by *ainult* '(if) only' in (52), which functions there as illocutionary adverb. Hence, these main clause usages of complementizers should be relegated to a more advanced stage of the insubordination process (see Evans 2007: 370, 374), where the formally subordinated clause is already conventionalized as main clause structure, and where the ellipted main clause cannot be restored, at least in a way preserving the illocutionary element in the subordinate clause.

Another group of occurrences of THAT-complementizers in main clauses concern cases where a superordinate perception predicate can be construed, and thus the discourse framing is evidential (see Laury & Seppänen 2008). This type of evidential insubordination is also typologically common (Evans 2007: 368,

16 See (Vallauri 2004) about the "denotational lightness" of the conceivable main clause.

394); cf. the Finnish sentence in (53), which could be based on visual evidence, report or inference.[17]

(53) ***Että*** *onkin* *kaunis* *paita!*
that be.PRS.3SG:CL beautiful.NOM shirt.NOM
'(And so) it is a beautiful shirt!'

A more idiosyncratic use of a complementizer in main clause is exemplified by (54) from Finnish, where *että* occurs in formulaic exclamative with the interjection *voi* 'oh'.

(54) *Voi* ***että*** *sun* *piti* *sanoa* *se!*
oh **that** you:GEN must:PST.3SG say:*A/DA* this.NOM
'Oh my, you had to go and say this!' (lit. Oh that you had to say this!).

Voi että (without the complement) is a common exclamation in colloquial Finnish, which expresses an unpleasant surprise. The complementizer *että* seems to function as a metalinguistic sign of complementation and the expression *voi että* seems to mark an exclamation over any possible complement, or over some generalized notion of clausal complementation.

The other complementizer types occur in main clauses as well. As mentioned, Q-complementizers occur in main clauses: they are formally identical with direct question markers. As for SIM-complementizers, the similative adverbs occur in independent clauses in which the verb is in the Conditional, although such clauses could be alternatively analysed as comparative adverbial clauses with omitted standard of comparison: consider (50) above without the matrix clause 'The prime minister acts'.

4.2 Complementizer omission

According to Boye, van Lier and Brink (2015), there is a tendency that omissible complementizers are epistemically neutral, in other words they do not contribute to the evaluation of the complement proposition. This cross-linguistic tendency is corroborated by Estonian, Finnish and Karelian, where only the general (semantically neutral) complementizers can be omitted; witness (55) from Estonian, where the complementizer *et* can be dropped.

17 See also (Keevallik 2000) on cases where the Estonian *et* functions as a synopsis marker.

(55) *Ma* *loodan* [***[et]*** *sa* *tuled* *ka*].
I.NOM hope:PRS.1SG **that** you.NOM come:PRS.2SG too
'I hope (that) you are coming too.'

Such sentences can be analyzed as juxtaposition of two clauses, that is, in terms of parataxis.[18]

The omissibility of the general complementizer does not depend purely on its semantics (resp. semantic neutrality), but is contingent on structural factors as well. Korhonen (1993: 116) noted that sentences like (55) are not acceptable in Finnish if the subject pronoun of the complement clause is covert[19], and they sound better if the matrix verb is in first or second person. But the Finnic languages behave differently in this respect: the first restriction for instance does not seem to be operative in Estonian (cf. ^{ok}*Loodan tuled ka*). Other factors operative in complementizer omission in Finnic include the relative order of the main and complement clause (only postposed complements allow for omission) and the category of the CTP: omission is never possible in complement clauses of nouns, cf. (56a) and (56b), or predicative adjectives, cf. (57a) and (57b), all from Finnish.

(56) a. *Pidän* *ajatuksesta,* [***että*** *kaikki* *vaikuttaa* *kaikkeen*].
like:PRS.1SG idea:ELA **that** everything influence:PRS.3SG everything:ILL
'I like the idea that everything has impact on everything.'
b. **Pidän* *ajatuksesta,* *kaikki* *vaikuttaa* *kaikkeen.*
like:PRS.1SG idea:ELA everything influence:PRS.3SG everything:ILL
Intended: 'I like the idea that everything has impact on everything.'

(57) a. *On* *hyvä,* [***että*** *hän* *liittyy* *meihin*].
be.PRS.3SG good **that** s/he join:PRS.3SG we:ILL
'It is good that s/he joins us.'
b. **On* *hyvä,* *hän* *liittyy* *meihin.*
be.PRS.3SG good s/he join:PRS.3SG we:ILL
Intended: 'It is good that s/he joins us.'

A similar class of complementizer-less sentences are those containing combinations of a frequent CTP and a first person singular pronoun (e.g. 'I think') that are

18 Tests for distinguishing between clause combining types (subordination, coordination and cosubordination), like the illocutionary scope test (Bickel 2010), are not easily applicable to sentences with omitted complementizers. For example, if the sentence in (55) is questioned with a polar interrogative, the complementizer becomes obligatory. As the interrogative is not compatible with complementizer omission, its illocutionary scope (conjunct, disjunct or else) cannot tell us with which clause combining type we are dealing.

19 Similar restrictions have been observed in other languages; compare the obligatoriness of the subject as a factor conditioning complementizer omission in Icelandic (Nordström & Boye, this volume).

being grammaticalized into parenthetical constructions occurring in the middle of discontinuous clauses; consider (58) from Estonian.

(58) *Ta sõidab, ma arvan [**et**], päris hästi*].
s/he drive:PRS.3SG I guess:PRS.1SG **that** quite well
'S/he drives, I guess, quite well.'

According to van Lier and Boye (2009), complementizer omission is related to grammaticalization of complement taking predicates. In colloquial Estonian the complementizer *et* is often omitted in contexts like (58) (Hennoste 2004; Keevallik 2003: 75–98; Keevallik 2010). The grammaticalization of subject-verb sequences, like *ma arvan* 'I think' or *ma usun* 'I believe', into parentheticals is manifested by their defectiveness; they are not grammatical under negation, see (59), they do not occur in certain tenses (as perfect or pluperfect), see (60), and they are usually incompatible with adverbial modifiers, see (61).

(59) **Ta sõidab, ma ei arva, (et) päris hästi.*
s/he drive:PRS.3SG I NEG guess.CN that quite good
'?S/he drives, I don't think, quite well.'

(60) **Ta sõidab, ma olen arvanud (et), päris hästi.*
s/he drive:PRS.3SG I be:PRS.1SG guess:PST.PTCP that quite good
'?S/he drives, I have guessed, quite well.'

(61) ?*Ta sõidab, ma arvasin eile (et), päris hästi.*
s/he drive:PRS.3SG I guess:PST:1SG yesterday that quite good
'?S/he drives, I thought yesterday, quite well.'

4.3 Complementizer combination

While complementizer omission may be seen as a case of underspecification of the interclausal relation, co-occurring complementizers may be seen as a case of over-specification of the dependency of clauses.

In the only frequently attested combination of two complementizers in Finnic Q-complementizers occur after THAT-complementizers. The semantic functions of such co-occurrences will be discussed in Section 5.4, but for the time being consider the Estonian example (62) with polar question marker occurring after the general complementizer.

(62) *Peeter küsis, [**et kas** Jaan suudles Marit*].
Peeter ask:PST.3SG [**that PQ** Jaan kiss:PST.3sg Mari:PART
'Peeter asked if Jaan kissed Mari.'

Essentially, this is the only possible relative order of co-occurring complementizers in Estonian, Finnish and Karelian: the general (or semantically neutral) complementizer always comes first and the second complementizer (in this case the polar marker) is within its scope (see also Korhonen 1993: 21 for Finnish).

Besides Q-complementizers, also SIM-complementizers co-occur with THAT-complementizers; cf. (63) from Olonets.

(63) *Īvan* *sanoi*, [***što*** ***buittaku*** *sinä* *djätid* *händy* *yksin* *koissa*].
Ivan say:PST.3SG **that** **as.if** you leave:PST.2SG he:PART alone home
'Ivan said that you (reportedly) left him alone at home.'
(EY and GV)

I am aware of only one combination of finite clause complementizers in Finnic, which does not contain the THAT-complementizer. This combination is however marginal, as it occurs only in colloquial and journalese Finnish (as well as in some Finnish dialects), and is considered to be calqued from Swedish (Tuomas Huumo, p.c.). Consider (64), where the conditional adverbializer *jos* 'if', which functions also as a complementizer (recall example [14]), is combined with the polar question clitic *=ko*/*=kö*.

(64) *En* *tiedä* [***jos=ko*** *hän* *tulee*].
NEG:1SG know.CN **if=PQ** s/he come:PRS.3SG
'I don't know if s/he is coming.'
(Korhonen 1993: 127)

This case of borrowing from Swedish to Finnish can be described as a replication-upon-identity. In Swedish, conditional protases and embedded polar questions are introduced by the same marker (*om*) (Nordström 2010: 173), whereas Finnish uses *jos* for protases and *=ko*/*=kö* for polar questions. The Finnish response to the Swedish pattern of identity has been to combine the two markers into a single complex complementizer.

5 Semantic functions of complementation markers

In this section I will pinpoint semantic differences between complementation markers by comparing sentences which form minimal pairs.

5.1 The contrast between *object of perception* and *knowledge acquired*

The contrast between object of perception and knowledge acquired is encoded in terms of the distinction between state-of-affairs and propositional content (see Boye 2010a; Dik & Hengeveld 1991). This contrast is expressed in Finnic by the contrast between participles and THAT-complementizers; see (65) from Finnish and (66) from Estonian.

(65) a. *Näin* *Jannen* ***tulevan*** *kaupasta.*
see:PST:1SG Janne:GEN **come:PRS.PTCP:GEN** shop:ELA
'I saw Janne coming from the shop.'
b. *Näin,* [***että*** *Janne* *tulee* *kaupasta*].
see:PST:1SG **that** Janne.NOM come:PRS.3SG shop:ELA
'I saw that Janne was coming from the shop.'

(66) a. *Nägin* *Jaani* *poest* ***tulevat.***
see:PST:1SG Jaan:PART shop:ELA **come:PRS.PTCP:PART**
'I saw Jaan coming from the shop.'
b. *Nägin,* [***et*** *Jaan* *tuleb* *poest*].[20]
see:PST:1SG **that** Jaan.NOM come:PRS.3SG shop:ELA
'I saw that Jaan was coming from the shop.'

The sentences in (a) express that someone perceives someone's coming, whereas the sentences in (b) express that someone acquires knowledge of the fact that someone is coming. Given the premise that objects of perception are non-propositions while knowledge acquired has propositional status, it is easy to empirically back up the assumption that the contrast between object of perception and knowledge acquired is at stake here. Consider for example the unavailability of participles as opposed to the availability of THAT-complementizers to epistemic qualification after perception verbs; see (67a) and (67b) from Estonian.

(67) a. **Jaan* *nägi* *Marit* *ilmselt* *lugevat.*
Jaan see:PST.3SG Mari:PART obviously read:PRS.PTCP:PART
'*Jaan saw Mari obviously reading.'
b. *Jaan* *nägi,* [***et*** *Mari* *ilmselt* *loeb*].
Jaan see:PST.3SG **that** Mari.NOM obviously read:PRS.3SG
'Jaan saw that Mari was obviously reading.'

20 Note that (66a) and (66b) have different word orders in their complements. This is another piece of evidence for coercion between the complementation marker and the complement clause; recall the analysis of examples (19) and (20) and consult Harder (this volume) for a general discussion of coercion in complement structures.

As only propositions have truth-value, only they are open for epistemic judgment. (67b) could be interpreted as Jaan seeing some signs of Mari reading, i.e. as mental perception, rather than immediate perception.[21]

Another test adhering to this involves questioning the complement: only propositional complements can be questioned (Dik & Hengeveld 1991: 233; Boye 2010a); cf. (68), also from Estonian.

(68) **Jaan* *nägi* *Marit* ***lugevat***. – *Kuidas?*
Jaan see:PST.3SG Mari:PART **read:PRS.PTCP:PART** how
'*Jaan saw Mari reading. – How?'

(68) is acceptable only if the question concerns the act of perception, not if it concerns the activity described in the complement. This shows that the participle is not truth-functional here.

These remarks are valid for the Inessive form of the *MA*-infinitive as well. This form has progressive meaning and is used to describe a perceived situation, which is simultaneous with the situation described in the main clause; cf. (69)–(71) from Estonian.

(69) *Jaan* *nägi* *Marit* ***luge-mas***.
Jaan see:PST.3SG Mari:PART **read-*MA*:INE**
'Jaan saw Mari reading.'

(70) **Jaan* *nägi* *Marit* ***luge-mas***. – *Kuidas?*
Jaan see:PST.3SG Mari:PART **read-*MA*:INE** how
'*Jaan saw Mari reading. – How?'

(71) *Jaan* *nägi* *Marit* *ilmselt* ***luge-mas***.
Jaan see:PST.3SG Mari:PART obviously **come:*MA*.INE**
'Jaan saw Mari(,) obviously reading.'

With this construction too, questions are possible, but only if they address the fact of perception and not the activity described in the complement; see (70). On the other hand, example (71) seems to pose a problem for the claim that object-of-perception complements do not allow epistemic qualification. There is no doubt that the complement of the verb 'see' in (71) expresses object of perception. It seems, however, that this is a quasi-problem. The object perceived in this sentence is not [*Mari's reading*], but [*Mari*], *who seemed to be reading*. In other words, the difference between (67a) and (71) is that in (67a) Mari's reading is in the immediate scope of the verb 'see', whereas in (71) Mari's reading is not in the immediate

21 Cross-linguistically perception verbs like 'see' or 'hear' often have secondary uses as markers of indirect evidentiality (cf. Boye & Harder 2007; Boye 2010a).

scope of the verb 'see'. Sands (2000: 263–264) points out that the Finnish participle expresses "a picture in mind" which as such can be an object of immediate perception, whereas the Inessive form of the *MA*-infinitive expresses a continuing process that is not immediately perceived, but rather observed for a longer period of time (see also Sands 2000: 251). As a result the achievement verb 'see' receives the narrows scope reading in (71). A proof of this is the impossibility of (72) with the atelic verb 'watch', which takes the Inessive form of the *MA*-infinitive in its immediate scope.

(72) ***Jaan** *vaatas* *Marit* *ilmselt* ***luge-mas***.
Jaan watch:PST.3SG Mari:PART obviously **read-*MA*.INE**
'[?]Jaan watched Mari(,) obviously reading.'

5.2 Epistemological contrast

A semantic contrast similar to the constrast between English *that* and *if* in the sentences *He knows that Maria will come* and *He knows if Maria will come* (see, for instance, van Lier & Boye 2009; Nordström 2010: 91; Nordström & Boye, this volume) is expressed also in Finnic. Consider the opposition between the THAT-complementizer and the polar Q-complementizer in Finnish.

(73) a. *Mari* *ei* *tiedä,* [***että*** *äitinsä* *on* *elossa*].
Mari NEG.3SG know.CN **that** mother.NOM:3POSS be.PRS.3SG alive
'Mari doesn't know that her mother is alive.'
b. *Mari* *ei* *tiedä,* [*on*=***ko*** *äitinsä* *elossa*].
Mari NEG.3SG know.CN be.PRS.3SG=**PQ** mother.NOM:3SG.POSS alive
'Mari doesn't know if her mother is alive.'

The difference between sentence (a) and sentence (b) is that (a) implies that the speaker knows that Mari's mother is alive, whereas (b) does not imply that s/he knows whether Mari's mother is alive or not. In other words, these sentences exemplify a contrast related to speaker's knowledge which can be labelled 'epistemological contrast'.[22]

22 I concede that this is a peculiar application of the term 'epistemological'; in philosophy this notion usually concerns the concepts of 'knowing that', 'knowing how' and acquaintance-knowledge (e.g. Ryle 1951: 25–61), rather than 'knowing if'. One could probably analyze the distinction in (73) independently of knowledge (at least in the narrow sense of the term). Harder (this volume), for example, relegates the differences between (73a) and (73b) to the contrast between "potential fact" and "issue".

5.3 Epistemic contrast

Epistemic modality relates to speaker's certainty that the proposition described in the complement clause is true or not true. In (74a) from Finnish the speaker expresses higher degree of certainty that somebody is crying than in the sentence (74b).

(74) a. *Minusta* *tuntuu,* [***että*** *joku* *itkee* *yläkerroksella*].
I:ELA feel:PRS.3SG **that** somebody cry:IND.PRS.3SG upper.floor:ADE
'It seems that somebody is crying upstairs.'
b. *Minusta* *tuntuu,* [***ikäänkuin*** *joku* *itki-si* *yläkerroksella*].
I:ELA feel:PRS.3SG **as.if** somebody cry-COND.3SG upper.floor:ADE
'It feels like somebody is (were) crying upstairs.'

The switch between the general complementizer and the SIM-complementizer after perception verb serves to convey the different degree of certainty of the speaker.

I argued earlier that the contrast between the general complementizer and the participial complementation marker occurring after perception verbs accounts for the distinction between knowledge acquired and object of perception. Consider now the difference between these complementation markers after utterance verbs; cf. (75) from Estonian.

(75) a. *Ta* *ütles,* [***et*** *Jaan* *on* *juba* *läinud*].
s/he say:PST.3SG **that** Jaan.NOM be.PRS.3SG already go:PST.PTCP
'S/he said that Jaan has already gone.'
b. *Ta* *ütles* *Jaani* ***olevat*** *juba* *läinud.*
s/he say:PST.3SG Jaan:GEN **be:PRS.PTCP:PART** already go:PST.PTCP
'S/he said/claimed that Jaan has already gone.'

Sentence (75b) shows that after utterance predicates the participle expresses propositions (this is the case after predicates of propositional attitude as well; recall Table 4). More important here is that sentence (75a) is unmarked with respect to epistemic modality, whereas sentence (75b) contains an element of doubt. The Partitive case form of the present participle (marked with the suffix *vat*) is employed as marker of reported evidentiality and carries the conventionalized implicature of doubt. In complements of utterance verbs this implicature is foregrounded, because the reportativity is communicated already by the complement taking predicate and thus the epistemic uncertainty is the only contribution of the participial form to the meaning of the sentence.

Finally, consider the contrast between the Finnish polar Q-complementizer *=ko/=kö* and the COND-complementizer *jos* 'if'. In indirect questions they are mutually exchangeable, but with slightly different implicature; cf. (76a) and (76b).

(76) a. *Kysyin* *Marilta,* [*pääsee=**kö*** *hän* *tulemaan*].
ask:PST:1SG Mari:ABL can:PRS.3SG=**PQ** she.NOM come:*MA*.ILL
'I asked Mari if/whether she can come.'
(Korhonen 1993: 111)

b. *Kysyin* *Marilta* [***jos*** *hän* *pääsee* *tulemaan*].
ask:PST:1SG Mari:ABL **if** she.NOM can:PRS.3SG come:*MA*.ILL
'I asked Mari if she can come.'
(Korhonen 1993: 111)

According to Korhonen (1993: 111), (76a) is neutral in the sense that the speaker does not take a stance as to the probability that Mari will manage to come, while (76b) implies that the speaker has some doubts in her chances of coming.

Recalling the discussion in the last section, we can conclude that one and the same complementation marker can be marked with respect to one distinction and unmarked with respect to another distinction. While the Finnish polar Q-complementizer was marked with respect to the epistemological contrast defined in Section 5.2, it is unmarked (or neutral) with respect to epistemic modality.

5.4 Evidential contrast

The complement in example (74b) above is identified by the SIM-complementizer (with the lexical meaning 'as if, like') and the Conditional mood marking on the dependent verb. Importantly, the Conditional mood does not occur only after similative words, but also after THAT-complementizers; as already noted in Section 2.1.1.1, THAT-complementizers are the general, syntagmatically unrestricted complementizers. This means in turn that one can construct minimal pairs of sentences in Finnish and Estonian with a dependent verb in the Conditional mood, which differ only in terms of the lexical marker of complementation. The only difference between such sentences would be that one of them contains the equivalent of the English complementizer *that* and the other contains a similative word. Such minimal pairs should be accounted for in terms of evidentiality rather than epistemic modality; cf. (77a) and (77b) from Estonian.

(77) a. *Talle* *tundub,* [***et*** *Jaan* *tuleks* *[, kui* *me* *talle* *maksaksime]*].
s/he:ALL feel:PRS.3SG **that** Jaan come:COND.3SG if we him pay:COND:1PL
'S/he thinks that Jaan would come (if we paid him).'

b. *Talle* *tundub* [***nagu*** *Jaan* *tuleks*].
s/he:ALL feel:PRS.3SG **as.if** Jaan come:COND.3SG
'It seems to her/him as if Jaan is coming' (at the very moment of observation).

As can be seen from the translations, (77a) and (77b) differ in terms of source of information rather than in terms of certainty. While (77a) expresses logical infer-

ence, (77b) expresses direct sensory evidentiality, where the sensation remains unspecified.

The reading of the SIM-complementizer depends on the type of perception (or sensation) specified by the main verb. Witness the Estonian sentences in (78a)–(78b) containing a verb of auditory perception.

(78) a. *Ma kuulsin, [**et** Jaan tuleks [, kui me maksaksime talle]*].
I hear:PST:1SG **that** Jaan come:COND.3SG if we pay:COND:1PL him
'I heard that Jaan would come if we paid him.'
b. *Ma kuulsin [**nagu** Jaan tuleks*].
I hear:PST:1SG **as.if** Jaan come:COND.3SG
'I heard as if Jaan was coming' (I heard the sound of Jaan coming).

While (78a) encodes hearsay, the usual reading of (78b) is *direct auditory evidence* (although hearsay reading is also possible). Consider now (79a) and (79b) with a verb of visual perception.

(79) a. *Ma nägin, [**et** Jaan tuleks [, kui me maksaksime talle]*].
I see:PST:1SG **that** Jaan come:COND.3SG if we pay:COND:1PL him
'I saw that Jaan would come if we paid him.'
b. *Ma nägin [**nagu** Jaan tuleks*].
I see:PST:1SG **as.if** Jaan come:COND.3SG
'I saw like Jaan was coming' (I saw Jaan in a distance, probably coming).

Here (79a) encodes logical inference, while the reading of (79b) is *direct visual evidence*.

The differences observed in the minimal pairs (77), (78) and (79) can be resolved in terms of verb categorization: recall that immediate perception verbs often have secondary inferential or reportative uses.[23] What happens in the examples (a) is that the free Conditional (serving to express that the complement content is hypothetical) disambiguates the CTP as a predicate expressing indirect evidentiality. For instance, observing the conditions upon which certain activity could take place involves always a mental process.

Everything said about Finnish and Estonian in this section also applies to Karelian, with the only difference that the Karelian SIM-complementizers *niinku, buittaku* and *rounoku* do not require the Conditional mood marking on the complement verbs, but occur with the Indicative mood. In Northern Karelian, Olonets

23 Alternatively, one may postulate two different 'see' verbs or 'hear' verbs. I have chosen, however, to regard them as a single polysemous verb. According to Viberg (2008: 135), with such verbs "it is often difficult to draw a sharp line between perceptual and other meanings. [...] Even cases like *Peter saw a bird* involve categorization which often is immediate but sometimes can be experienced as a separate process".

and Lude one can construct minimal pairs of sentences equivalent to (77)–(79), where in the (a)-sentences the complement verb is in the Conditional whereas the complement verb in the (b)-sentences is in the Indicative.

The above discussion can be summarized by the observation that Finnic SIM-complementizers encode by default direct evidentiality, while THAT-complementizers encode indirect evidentiality. This is not surprising; there are many other languages in which complementizers mark or at least contribute to the marking of evidentiality distinctions (see Aikhenvald 2004: 120–123 for references). It is interesting to note, however, that in the face of SIM-complementizers we have direct evidentiality combined with low certainty and negative propositional attitude: the speaker conveys that she does not fully trust her senses.

One could be confident that although the complements in (77b), (78b) and (79b) occur after immediate perception verbs they express propositions and not state-of-affairs. This is indicated by the finiteness of the complement verb as well as by the complement being subjected to epistemic and evidential qualification. There are, however, cases in which it is more difficult to distinguish between the contrast 'knowledge acquired vs. object of perception' on the one hand, and 'indirect vs. direct evidentiality', on the other. Consider the interchangeability of THAT-complementizers and participles after perception verbs, which was argued in Section 5.1 to encode the contrast between knowledge acquired and object of perception; cf. examples (80a) and (80b) from Finnish.

(80) a. *Näen,* [***että*** *Maija* *menee*].
see.PRS:1SG **that** Maija.NOM go:PRS.3SG
'I see that Maija is going.'
(Sands 2000: 125)

b. *Näen* *Maijan* ***menevän***.
see.PRS:1SG Maija:GEN **go:PRS.PTCP:GEN**
'I see Maija going.'
(Sands 2000: 125)

Sands (2000: 125) argues that in the case of (80a) the speaker may be a witness of some, not necessarily direct evidence that Maija is leaving. In (80b), on the other hand, the speaker is a witness of the fact of Maija's leaving at the very moment this happens. In other words, the difference between (80a) and (80b) could be accounted for in terms of unmarked (direct or indirect) evidentiality versus direct (first-hand sensory) evidentiality.

This would imply identity between the notions of direct perception (object of perception) and direct evidence. Yet, Boye (2010b: 95) has argued that these notions refer to different phenomena: perceiving a state-of-affairs directly is not identical to having direct evidence for a proposition (i.e. information about the world), although the two may imply each other. That the minimal pair in (80)

exemplifies the distinction between acquired knowledge and perception of a state-of-affairs, and not evidentiality distinction, is attested by the fact that the complement content in (80b) cannot be challenged by an epistemic adverb or questioned which in turn means that the complement is not a proposition that could be evidentially qualified in the first place. The misconception in Sands analysis is due to the alternative readings of the CTP. While in (80a) the complement contents could be either directly observed or inferred, in (80b) the direct perception designated by the complement triggers only the direct evidentiality reading of the complement-taking verb.

5.5 Reported and quoted (indirect) speech

In addition to evidentiality proper, the Finnic complementizer system provides evidence for a contrast between *reported indirect speech* and something that can be provisionally labelled *indirect quotative*.[24] While reported indirect speech represents someone else's speech in one's own words, and direct speech is a quotation of someone else's utterance in its original form, the indirect quotative represents an intermediate way of encoding mismatch in authorship of utterances. Consider the Estonian examples in (81), which contain different structures used to encode a question addressed to the speaker (the origo in the sentence) by a man called *Jaan*.

(81) a. *Jaan küsis:* [***Kas*** *sa ei tulegi Tallinnasse?*]
Jaan ask:PST.3SG **PQ** you NEG come.CN:CL Tallinn:ILL
'Jaan asked: Aren't you ever going to come to Tallinn?'

b. *Jaan küsis,* [***kas*** *ma ei tulegi Tallinnasse?*]
Jaan ask:PST.3SG **PQ** I NEG come.CN:CL Tallinn:ILL
'Jaan asked, if I'm never going to come to Tallinn.'

c. *Jaan küsis,* [***et kas*** *sa / ma ei tulegi Tallinnasse?*]
Jaan ask:PST.3SG **that PQ** you I NEG come.CN:CL Tallinn:ILL
'Jaan asked [aren't you / am I] ever going to come to Tallinn.'

Example (81a) is a case of direct speech with subject in the second person, as produced by the author of the question. (81b) contains reported indirect speech with a deictic shift from second to first person in the complement clause (the indirect question). Finally, example (81c) presents an intriguing mixture of indirect and direct speech. On the one hand, this sentence contains two complementizers – *et* (THAT-complementizer) and *kas* (polar question complementizer), which

24 I adopt here Boye's (2012: 32) view that quotatives are not true evidentials.

means that the syntactic subordination, a distinctive feature of indirect speech, is not only expressed but overspecified (recall the discussion on complementizer combinations in Section 4.3). On the other hand, the grammaticality of both, the second and the first person pronoun in (81c), shows that this structure allows for both – preserving the original deictic orientation of the quotation, and for deictic shift characteristic for indirect speech (replacement of the second person marking with first person marking). The combination of these two complementizers may be regarded as a complex complementizer with specific meaning. One may consider the option with the second person pronoun *kas sa ei tulegi Tallinnasse* in (81c) as a quotation of a question and the combination *et* + *kas* as a quotative marker, which exhibits, however, formal characteristics of indirect speech. Upon this analysis, the occurrence of *kas* without *et* in (81b) encodes reported indirect speech.

Apart from being consistent with the intuitions of native speakers, this analysis receives support from previous research on the combination of THAT- and Q-complementizers in Estonian. Remmelg (2006: 66) observed that in contemporary written Estonian the complementizer combination in (81c) is often employed in order to avoid using direct speech punctuation.

5.6 Other semantic contrasts

Sands (2000: 145) analyses the Finnish examples in (80) in terms of volitional control.

(82) a. *Harkitsin* *jäisinkö* *Babyloniin* *opettelemaan*
consider:PST:1SG remain:COND:1SG:PQ Babylon:ILL learn:*MA*.ILL
tähtien ***lukemista***.
star:PL:GEN **read:ACNMLZ**
'I considered whether to remain in Babylon to learn the reading of the stars.'
(Sands 2000: 145)

b. *Harkitsin* *jäisinkö* *Babyloniin* *opettelemaan*
consider:PST:1SG remain:COND:1SG:PQ Babylon:ILL learn:*MA*.ILL
lukemaan *tähdet*.
read:*MA*.ILL star:PL.NOM
'I considered whether to remain in Babylon to learn to read the stars.'
(Sands 2000: 145)

The Illative form of the *MA*-infinitive of the verb 'read' in (82b) implies that the protagonist remained in Babylon "in order to learn how to read the stars". Here the subject has control over the learning process. With the action nominal in (82a), on the other hand, the protagonist learns how to read the stars "more through absorption rather than by actively setting out to learn" (Sands 2000: 145). In other

words, the Illative form of the *MA*-infinitive implies as a complementation marker higher degree of volitional control than the action nominalization. This could be due to the fact that the Finnic *MA*-infinitive has an implicit but specific actor argument, whereas the action nominal has a generic actor argument (cf. 'I stayed in Babylon to learn how *people* learn to read the stars.' as a paraphrase of [82a]). This implies mapping between specific actors and strong volitional control on the one hand, and generic actors and weak volition on the other.[25]

5.7 Complex semantic contrasts

So far I have discussed only examples of binary semantic contrasts. It remains to be resolved whether the Finnic languages testify for more complex contrasts. The clearest example of a tripartite contrast, which I am aware of, is presented in (83) from Estonian.

(83) a. *Sa* *saad* *teada* [***kui*** *ta* *tuleb* *linna*].
you get:PRS.2SG know:*A/DA* **if/when** s/he come:PRS.3SG town.ILL
'You will get to know when s/he comes to the town.'

b. *Sa* *saad* *teada,* [***et*** *ta* *tuleb* *linna*].
you get:PRS.2SG know:*A/DA* **that** s/he come:PRS.3SG town.ILL
'You will get to know that s/he is coming to the town.'

c. *Sa* *saad* *teada,* [***kas*** *ta* *tuleb* *linna*].
you get:PRS.2SG know:*A/DA* **PQ** s/he come:PRS.3SG town.ILL
'You will get to know if s/he is coming to the town.'

The sentences in (83) exemplify a tripartite contrast between the TEMP/COND-complementizer, the THAT-complementizer and the Q-complementizer. The dependent clause in (83a) can be argued to function as a temporal adjunct. Note, however, that this clause, headed by *kui*, is the only evident object of the verb 'know'. What the sentence means is 'You will get to know the fact of him/her coming to the town'. At the same time, the lexical semantics of the temporal subordinator *kui* have impact on the meaning of the sentence and its temporal interpretation is specific: the fact of someone being informed about someone's arrival in the town is construed as (roughly) coinciding with the arrival. In contrast, in (83b) and (83c) the fact of someone being informed about someone's arrival is not related to a specific point in time and the temporal interpretation of these sen-

25 This mapping could be compared to the typologically frequent association between impersonals and non-volitional contexts.

tences is non-specific.[26] What concerns the complement proposition, in (83a) and (83b) it is considered as factual (i.e. it is implied that s/he will come to the town), while (83c) does not imply whether the agent of the complement clause will come or not (recall the epistemological contrast discussed in Section 5.2).

As can be seen, the first complementizer enters into a time reference opposition with the second and the third, whereas the third complementizer enters into a factuality opposition with the first and the second; this is illustrated in Table 5.

Table 5: Tripartite contrast in Estonian: TEMP/COND-complementizer, THAT-complementizer and Q-complementizer

	time reference	**factuality**
TEMP/COND-complementizer	specific	implied
THAT-complementizer	unspecified	implied
polar Q-complementizer	unspecified	not implied

The distribution in the table shows again that THAT-complementizer is the neutral (or unmarked) member of the complementizer system. Clearly, specific time reference is marked with respect to the lack of specific time reference. Likewise, information which is communicated but not implied is expected to be formally more complex than implied (factual) information. This means that the TEMP/COND-complementizer and the polar Q-complementizer are both marked with respect to one of the semantic features under concern, whereas the THAT-complementizer is unmarked with respect to both features.

5.8 Semantic functions of complementation markers: summary

Table 6 presents a condensed picture of the semantics landscape. I have included Karelian to the last column in those cases where the given contrast is attested in at least one of the Karelian dialects studied here.

26 One might be tempted to account for this contrast in terms of the temporal-aspectual distinction between 'episodic' and 'generic'. Note, however, that while (83a) has a prototypical episodic reading, (83b) and (83c) are not generic in the sense found in the literature on tense and aspect (for example, the propositional content of these sentences does not hold true during the time of speech; see Dahl 1995: 424).

Table 6: Summary of the semantic contrasts of complementation markers

semantic contrast	types of complementation markers involved	languages
knowledge acquired vs. object of perception	*THAT-complementizer* vs. *participle*	Estonian, Finnish, Karelian
	THAT-complementizer vs. *Inessive form of MA-infinitive*	Estonian, Finnish, Karelian
epistemological contrast	*THAT-complementizer* vs. *polar Q-complementizer*	Estonian, Finnish, Karelian
	TEMP/COND-complementizer vs. *polar Q-complementizer*	Estonian, Finish, Karelian
epistemic contrast	*THAT-complementizer* vs. *SIM-complementizer*	Estonian, Finnish, Karelian
	that-complementizer vs. *participle*	Estonian, Finnish, Karelian
	polar Q-complementizer vs. *COND-complementizer*	Finnish
evidential contrast	*THAT-complementizer* (+ Conditional mood) vs. *SIM-complementizer*	Estonian, Finnish, Karelian
reported vs. quoted (indirect) speech	*Q-complementizer* vs. *THAT-complementizer & Q-complementizer*	Estonian, Finnish, Karelian
volitional control	*Illative form of MA-infinitive* vs. *action nominal*	Estonian, Finnish, Karelian
temporality	*THAT-complementizer* vs. *TEMP/COND-complementizer*	Estonian, Finnish, Karelian
	polar Q-complementizer vs. *TEMP/COND-complementizer*	Estonian, Finnish, Karelian

The distribution in the table shows that some pairs of complementation markers engage in the marking of different semantic contrast. This is a potential source of misconception and one of the disadvantages of the onomasiological approach to complementation marker semantics assumed in this study. The first case in point is provided by the THAT-complementizer and the participle, which both encode the contrast 'knowledge acquired vs. object of perception' and the epistemic contrast (different degrees of certainty). Consider however the different complement taking predicate: while these two complementation markers convey the contrast between knowledge acquired and object of perception after perception verbs, they convey the epistemic contrast after verbs of utterance and report; compare examples (65)–(66) with example (75).

Likewise, the THAT-complementizer and the SIM-complementizer are used not only to encode an epistemic distinction, but also to encode the distinction between indirect and direct evidentiality. Note, however, that in the second distinction, the general complementizer is followed by a Conditional mood on the complement verb while in the first distinction it is followed by the Indicative.

It could be assumed that this syncretism in the linguistic coding of propositionality, epistemicity and evidentiality reflects the strong implicational relation between notions such as direct perception (of a state-of-affairs) and sensory evidentiality or between uncertainty and indirect evidentiality (see Boye 2012 for discussion).

The third and the last pair of complementation markers involved in the marking of different semantic contrasts consist of the TEMP/COND-complementizer and the polar Q-complementizer. As can be seen from the table, this pair encodes an epistemological and a temporality contrast. Note, however, that in (83) the TEMP/COND-complementizer was the marked member and the polar Q-complementizer the unmarked member of the temporal opposition, whereas in the factuality opposition their relative markedness is reverse.

6 Diachronic sources of complementizers

The question of how the semantic functions of Finnic complementation markers exactly developped has to wait for further research. There are no studies in the field dealing with the development of semantic contrasts between complementation markers. Therefore, I will focus in this section on the development of certain forms into complementation markers and only occasionally will include remarks about semantics.

The finite and non-finite complementation markers will be dealt with in separate sections, but first I will discuss the relative distribution of these two types at earlier periods of language history. Such discussion will help to determine the starting point for the development of the Finnic complementation marker system(s).

6.1 Relative distribution of the finite and non-finite complementation markers in the history of Finno-Ugric

According to the prevailing view in Finno-Ugric studies (see Karelson 1959: 70 for references), at a certain earlier stage of Finno-Ugric, clauses were linked either by

mere juxtaposition (cf. example [84] from Liv) or by non-finite verb morphology (cf. example [85] from Finnish, with a participial complement).

(84) *Ma näis, pois itkūb*
I see:PST.1SG boy.NOM cry:PRS.3SG
'I saw, the boy is crying.'
(Karelson 1959: 70)

(85) *Näin pojan itkevän.*
see:PST:1SG boy:GEN cry:PRS.PTCP:GEN
'I saw the boy crying.'

The preeminence of juxtaposition and non-finite strategies compared to finite subordination with conjunctions is easily deduced from the relatively recent emergence of conjunctions in Finno-Ugric. In fact, most conjunctions in the Finnic languages are loans from Indo-European. Examples include Liv *un* 'and' (< Latvian < German; see Kettunen 1938: 452), Estonian and Finnish *ja* 'and' (< Germanic), and Eastern Finnic *a* 'but', *da* 'and; but; though', *što* 'that', *hoť* 'though' and *jesli* 'if' (< Russian) (see Karelson 1959: 66 for more examples).

The importance of non-finite subordination (and complementation in particular) can be deduced from the fact that easternmost Finno-Ugric languages are nowadays characterized by prevalence of non-finite over finite subordinate clauses. These languages came into contact with Indo-European later than Finnic and therefore are believed to have preserved the Proto-Finno-Ugric syntax to a greater extent. According to certain evolutionist view, which goes back at least to Ravila (1961: 113–119), and which is still dominant in Finno-Ugric linguistics, Proto-Finno-Ugric was characterised by a lack of finite verb forms and, respectively, by a lack of finite clauses. Instead of *the bird flies* ancient Finno-Ugrians supposedly used a construction with a nominal form of the verb such as *the bird flying*. According to the reconstructed scenario, after the emergence of finite verb forms non-finite verb forms were relegated to the expression of clausal subordination (as in *I heard him coming*), and in eastern Finno-Ugric this is still the major means for marking clausal subordination.[27]

If we accept the view that the proto-language was characterized by non-finite complement clauses, and that the construction with a complementizer and a finite complement verb is later development, we have to cope with the following problem. Considering the fact that in contemporary Finnic state-of-affairs is basically a semantic property of non-finite complements and propositional content

27 A similar scenario for Indo-European (and Baltic in particular) was suggested by Ambrazas (1990: 236–237), who showed that in the earliest texts from Baltic languages non-finite subordination prevailed over finite subordination using conjunctions.

of finite complements, the question arises of how complement propositions were expressed in Proto-Finnic, for which no finite complements are reconstructed. The only solution would be to assume that non-finite complements have been formerly employed to also express propositions, much like they are currently employed in the easternmost Finno-Ugric languages, and partly also in Finnic (recall the occurrence of participles after predicates of propositional attitude and knowledge discussed in Sections 3 and 4).

6.2 The development of non-finite complementation markers

Despite of the congenitally nominal (non-finite) character of the ancient Finno-Ugric clause, scholars have been looking to identify the bridging contexts for the reanalysis of participles and infinitives into complementation marking devices.

A popular scenario for the development of participles into complementation markers assumes reanalysis of the argument structure of a construction with the participle (Hakulinen 2000: 571–572). In particular, a participial modifier has acquired the status of complement predicate by taking over an argument of the main verb. Consider the sentences in (86) and (87).

(86) *Kuulin* *käe-n* *kukku-va-n.*
hear:PST:1SG cuckoo-ACC call-PRS.PTCP-ACC
'I heard the cuckoo's calling.'
(Hakulinen 2000: 571–572)

(87) *Kuulin* *käe-n* *kukku-van.*
hear:PST.1SG cuckoo-GEN call-PRS.PTCP:GEN(=COMP)
'I heard the cuckoo calling.'
(Hakulinen 2000: 571–572)

In (86) *käen* (nominative *käki*) is the object of the verb *kuulin* 'I heard' which can be seen from its Accusative case marking. The participial form *kukkuvan*, on the other hand, functions as adjective complement of *käki* 'cuckoo'. In terms of phrase structure (86) is parallel to (88).

(88) *Veistin* *puiko-n* *liian* *lyhye-n.*
carve:PST:1SG stick-ACC too short-ACC
'I carved the stick too short.'
(Hakulinen 2000: 571–572)

In many grammatical contexts, however, the Accusative is homonymous with the Genitive, which has led to reanalysis of the argument structure, exemplified in (87). In this example *käen* is no longer an object of the finite verb, but a subject of

the participle in non-finite predication taking a Genitive subject. In other words, we observe here promotion of the participle from the status of adjective complement to the status of predicate of the complement clause, which takes arguments on its own.

As for the infinitives, the reconstruction goes back to their nominal origins.[28] The *A/DA*-infinitive originally has been a deverbal noun marked by the lative case suffix *-k*, which later disappeared. Its initial meaning as lative deverbal noun has been to express direction or purpose (Hakulinen 2000). In contemporary Finnish it can still be sometimes exchanged for the purpose marking action nominal in Illative or Translative case; cf. the examples in (89).

(89) a. *Anna* *leipää* *pojan* *syödä.*
give.IMP.2SG bread:PART boy:GEN eat:*A/DA*(*-LAT)
'Give some bread to the boy to eat.'
(Saukkonen 1965: 209)

b. *Anna* *leipää* *pojan* *syömiseen / syömiseksi.*
give.IMP.2SG bread:PART boy:GEN eat:ACNMLZ:ILL eat:ACNMLZ:TRANS
'Give some bread to the boy to eat / for eating.'
(Saukkonen 1965: 209)

6.3 The development of finite clause complementizers

The vast majority of the canonical complementizers in Finnic are subordinating conjunctions. Most of them descend from lexical words and can be classified according to the word class of their source. Their semantics usually mirror their lexical semantics from the pre-complementizer stage. During their development into complementizers, they have lost semantic substance and become more general in meaning, although this change has its natural boundaries: a language can have only one semantically neutral complementizer, which means that only the winner is fully desemanticized.

A large number of complementizers derive from pronouns. There are two types of pronouns, which have given rise to complementizers: interrogative pronouns and demonstrative pronouns. Consider the following example of the development of interrogative pronoun into complementizer. The interrogative pronoun stem *mi* was grammaticalized in Veps into a complementizer with the lexical meaning 'how, like; that'. Karelson (1959: 202) has suggested the following path of devel-

28 This seems to be frequent scenario cross-linguistically. According to Noonan (2007), in several languages infinitives are derived from nominal constructions; for example, many Indo-European infinitives are diachronically case-marked nominalizations.

opment: the starting point of the grammaticalization process was the Instructive (instrumental) form of this pronoun *min*, which was later shortened into *mi*. The initial meaning of this inflectional form was 'with what; by what means'. This meaning evolved then to the semantically less complex 'how', which ultimately developed to 'like', conforming thus to the Russian polysemous word *kak*, which means both 'how' and 'like'. Finally, this word was desemanticized into a general complementizer (see Karelson 1959: 203 for examples). Another old Finno-Ugric interrogative stem that has served as a source of many complementizers in the contemporary languages is *ko-/ku-*. Its descendants are the Finnish TEMP-complementizer *kun*, the Estonian TEMP/COND-complementizer *kui* 'when; if' and the Olonets general complementizer *ku* (see Karelson 1959: 145–162).

Examples of complementizers originating in inflected forms of demonstrative pronouns include the THAT-complementizer *et(tä)*, which is attested in the majority of Finnic languages. Historically, it consists of the demonstrative pronoun stem *e-* and the Partitive or Abessive case marker (Karelson 1959: 100–105). The functional development of this complementizer is identical to the development of the English *that* or German *das(s)* into complementizers (see Laury & Seppänen 2008: 172–173 for a detailed description of the developments in Finnic).

There are fewer cases in which complementizers have adverbs as their sources. Many inflected forms of pronouns have been lexicalized into adverbs. In some cases a lexical adverb is the intermediate stage between an inflected form of a pronoun and a complementizer.[29] Examples include the Estonian *kuni*, Finnish *kunnes* or Olonets *kunne/kunna*, which are originally Terminative case forms of the pronominal *ku*-stem, but are employed as adverbs (all meaning 'until'), and, as example (15) suggests, marginally also as complementizers.

Finally, one should consider the lexicalization of multi-word expressions as a source of complementizers. Typical examples of lexicalized compound complementizers are the Estonian similative words *justkui* (*just* 'just' + *kui* 'when; if'), *justnagu* (*just* 'just' + *nagu* 'like') and *otsekui* (*otse* 'straight' + *kui* 'when; if'), all meaning approximately 'as if; like; allegedly'. Their Finnish translation equivalent *ikäänkuin* reveals even more complex structure: *ikään* is an Illative case form of the noun *ikä* 'age' and *kuin* is 'like' (Karelson 1959: 336).

29 Note that lexicalization (as storage) can be a part of the process of grammaticalization (see Rostila 2006).

7 Conclusions

The Finnic languages have two parallel systems of complementation markers: a) finite clause complementizers (mostly lexical subordinators), and b) non-finite verb forms with their constructional properties. The co-existence of these two systems in the modern languages reflects the general antagonism between finite and non-finite subordination in Finno-Ugric. The non-finite subordination is losing ground to the finite one, stubbornly keeping its positions in the standard written languages. It can be predicted that high literary registers will be the last resort of the non-finite complementation markers.

While THAT-complementizers can be seen as neutral with respect to the choice of complement taking predicate, immediate perception predicates is the neutral type of CTP with respect to the choice of complementation marker.

All forms functioning as finite clause complementizers occur in main clauses, and nearly all occur in other types of dependent clauses.

Only semantically neutral complementizers can be omitted; their omissibility is restrained, however, by various structural factors. Complementizer combination, on the other hand, is not only possible in Finnic, but may have semantic import on its own.

The study identified several semantic functions of complementation markers in Finnic. It was demonstrated that complementation markers are used to code the conceptual complexity of the complement (proposition vs. state-of-affairs), to express availability of facts (knowledge), and to express distinctions related to modality, evidentiality, volition and temporality. One and the same complementation marker may participate in the coding of different semantic contrasts.

Acknowledgment

This investigation was carried out during my research stay at the University of Regensburg, sponsored by Alexander von Humboldt Foundation and the Graduate School for East and Southeast European Studies. Further support was provided by grant no. 7476 of the Estonian Science Foundation and project no. 10–202 of the program *The Estonian Language and Cultural Memory* of the Estonian Ministry of Education and Science. I am particularly grateful to Sandra Birzer, Kasper Boye, Mai Frick, Björn Hansen, Tuomas Huumo and to my colleagues from the research group *Morphosyntactic structure and development of Estonian* from the University of Tartu for stimulating discussion and valuable hints, and to the Department of Slavic Philology at the University of Regensburg for their

assistance in practical matters. I am also grateful to Maria Urbanovich who organized my stay in Jyškyjärvi and to my Karelian consultants Maria Kopulova, Aino Tarasova, Vladimir Parfent'ev, Antonina Parfent'eva, Zinaida Vasilieva, Galina Stepanova, Lidia Sokolova, Lidia Rugoeva and Zoya Nefantieva from Jyškyjärvi, to Ekaterina Efimova and Galina Vasilieva from Kindasovo, and Zinaida Lapeka, Vasiliĭ Savel'ev, Anja Savel'eva and Mikhail Titov from Priäžä.

References

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.

Ambrazas, Vytautas. 1990. *Sravniteľnyĭ sintaksis prichastiĭ baltiĭskikh yazykov – Vergleichende Syntax der baltischen Partizipien* [Comparable syntax of Baltic participles]. Vilnius: Mokslas.

Bickel, Balthazar. 2010. Capturing particulars and universals in clause linkage: a multivariate analysis. In Isabelle Bril (ed.), *Clause-hierarchy and clause-linking: The syntax and pragmatics interface*, 51–101. Amsterdam & Philadelphia: John Benjamins.

Boye, Kasper. 2010a. Reference and clausal perception-verb complements. *Linguistics* 48(2). 391–430.

Boye, Kasper. 2010b. Evidence for what: Evidentiality and scope. *Sprachtypologie und Universalienforschung* 63(4). 290–307.

Boye, Kasper. 2012. *Epistemic meaning. A crosslinguistic and functional-cognitive study* (Empirical Approaches to Language Typology 43). Berlin: De Gruyter Mouton.

Boye, Kasper & Peter Harder. 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in Language* 31(3). 569–606.

Boye, Kasper, Eva van Lier & Eva Theilgaard Brink. 2015. Epistemic complementiers: a cross-linguistically survey. *Language Sciences* 51. 1–17.

Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.

Dahl, Östen. 1995. The marking of the episodic/generic distinction in tense-aspect systems. In Gregory N. Carlson & Jeffrey F. Pelletier (eds.), *The generic book*, 412–425. Chicago: University of Chicago Press.

Dik, Simon C. & Kees Hengeveld. 1991. The hierarchical structure of the clause and the typology of perception-verb complements. *Linguistics* 29(2). 231–259.

Dixon, Robert M. W. 2006. Complement clauses and complementation strategies in typological perspective. In Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation: A cross-linguistic typology*, 1–48. Oxford: Oxford University Press.

EKG II = Erelt, Mati, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael & Silvi Vare. 1993. *Eesti keele grammatika II. Süntaks. Lisa: kiri* [Estonian grammar II. Syntax. Appendix: Written language]. Tallinn: Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut.

Evans, Nickolas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness. Theoretical and empirical foundations*, 366–431. Oxford: Oxford University Press.

Hakulinen, Lauri. 2000. *Suomen kielen rakenne ja kehitys* [The structure and development of the Finnish language]. Helsinki: Helsingin yliopiston suomen kielen laitos.

Harder, Peter. This volume. Complementizers and the complementizing function. Structure and pragmatics in a panchronic perspective.
Hennoste, Tiit. 2004. *Et*-komplementlause kesksete põhiverbide funktsioonid eestikeelses vestluses [The functions of the common matrix verbs of *et*-complements in Estonian conversation]. *Keel ja Kirjandus* 2004(7). 504–529; 2004(8). 590–609.
Herlin, Ilona. 1998. Suomen *kun* [The Finnish *kun*] (Suomalaisen Kirjallisuuden Toimituksia 712). Helsinki: Suomalaisen Kirjallisuuden Seura.
Holvoet, Axel. This volume. Semantic functions of complementiers in Baltic.
Karelson, Rudolf. 1959. *Sidesõnad läänemeresoome keeltes* [Conjunctions in Finnic languages]. University of Tartu: Doctoral thesis, Manuscript.
Karttunen, Lauri. 1971. Some observations on factivity. *Papers in Linguistics* 4(1). 55–69.
Keevallik, Leelo. 2000. Keelendid *et* ja *nii et* vestluses [The linguistic units *et* and *nii et* in conversation]. *Keel ja Kirjandus* 43(5). 344–358.
Keevallik, Leelo. 2003. *From interaction to grammar. Estonian finite verb forms in conversation* (Acta Universitatis Upsaliensis. Studia Uralica Upsaliensia 34). Uppsala: Uppsala University.
Keevallik, Leelo. 2010. Clauses emerging as epistemic adverbs in Estonian conversation. *Linguistica Uralica* 46(2). 81–101.
Kettunen, Lauri. 1938. *Livisches Wörterbuch mit grammatischer Einleitung* (Lexica Societatis Fenno-Ugricae 5). Helsinki: Druckerei-A.G. der finnischen Literaturgesellschaft.
Korhonen, Riitta. 1993. *Buts about conjunctions. A syntactic study of conjunctions expressions in Finnish*. Helsinki: Suomalaisen Kirjallisuuden Seura.
Laanest, Arvo. 1982. *Einführung in die ostseefinnischen Sprachen*. Hamburg: Helmut Buske.
Laury, Ritva & Eeva-Leena Seppänen. 2008. Clause combining, interaction, evidentiality, participation structure, and the conjunction-particle continuum. The Finnish *että*. In Ritva Laury (ed.), *Crosslinguistic studies in clause combining. The multifunctionality of conjunctions* (Typological Studies in Language 80), 153–178. Amsterdam & Philadelphia: John Benjamins.
Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse* (Typological Studies in Language 18), 181–226. Amsterdam: John Benjamins.
Letuchiy, Aleksander. 2008. Sravniteľnye konstruktsii, irrealis i evidenciaľnosť [Comparative constructions, irrealis and evidentiality]. In Björn Wiemer & Vladimir A. Plungjan (eds.), *Lexikalische Evidenzialitätsmarker in slavischen Sprachen* (Wiener Slawistischer Almanach, Sonderband 72), 215–238. München: Otto Sagner.
Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
Markianova, Ljudmila. 1993. *Livvin murdehen morfolougii. Nominat da abusanat. Opetus-moniste* [Morphology of Olonets dialect. Nouns and function words. Teaching materials]. Petroskoi: Petroskoin valdivonüliopisto.
Metslang, Helle. 1999. Is the Estonian and Finnish conditional actually a conditional? *Estonian: Typological Studies III* (Tartu Ülikooli eesti keele õppetooli toimetised 11), 97–127. Tartu: University of Tartu.
Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. II: *Complex constructions*, 2nd edn., 52–150. Cambridge: Cambridge University Press.
Nordström, Jackie. 2010. *Modality and subordinators* (Studies in Language Companion Series 116). Amsterdam & Philadelphia: John Benjamins.

Nordström, Jackie & Kasper Boye. This volume. Complementizer semantics in the Germanic languages.
Palmer, Frank Robert. 2001. *Mood and modality*, 2nd edn. Cambridge: Cambridge University Press.
Ravila, Paavo. 1961. *Johdatus kielihistoriaan* [Introduction to Language History] (Tietolipas 3). Helsinki: Suomalaisen Kirjallisuuden Seura.
Remmelg, Raili. 2006. Et-*kõrvallaused tänapäeva eesti kirjakeeles* [*Et*-headed dependent clauses in contemporary Estonian]. University of Tartu: MA thesis, Manuscript.
Rostila, Jouni. 2006. Storage as a way to grammaticalization. *Constructions* 2006(1) www.constructions-online.de (accessed March 2012).
Ryle, Gilbert. 1951. *The concept of mind*. New York, Melbourne, Sidney & Cape Town: Hutchinson's University Library.
Sands, Kristina. 2000. *Complement clauses and grammatical relations in Finnish*. Australian National University: PhD thesis, Manuscript.
Saukkonen, Pauli. 1965. *Itämerensuomalaisten kielten tulossijainfinitiivirakenteiden historiaa I* [A history of lative infinitive constructions in Finnic languages I]. Helsinki: Suomalais-Ugrilainen Seura.
SKJ-LD = Makarov, Grigorij N. 1990. *Slovar' kareľskogo yazyka (livvikovskiĭ dialekt)* [Dictionary of Karelian. Olonets dialect]. Petrozavodsk: Karelija.
Thompson, Sandra A. 2002. 'Object complements' and conversation. *Studies in Language* 26(1). 125–164.
Turunen, Aimo. 1946. *Lyydiläismurteiden äännehistoria. I, Konsonantit* [Historical phonology of Lude dialects. I, consonants] (Mémoires de la Société Finno-ougrienne 99). Helsinki: Suomalais-Ugrilainen Seura.
Vallauri, Edoardo Lombardi. 2004. Grammaticalization of syntactic incompleteness: free conditionals in Italian and other languages. *SKY Journal of Linguistics* 17. 189–215.
van Lier, Eva & Kasper Boye. 2009. Epistemic complementizers: A cross-linguistic survey. Paper presented at the Eighth Biennial Conference of the Association for Linguistic Typology (ALT), Berkeley, 23–26 July.
Viberg, Åke. 2008. Swedish verbs of perception from a typological and contrastive perspective. In María de los Ángeles Gómez González, Lachlan Mackenzie & Elsa González Álvarez (eds.), *Languages and cultures in contrast: New directions in contrastive linguistics* (Pragmatics and Beyond New Series 175), 123–172. Amsterdam & Philadelphia: John Benjamins.
VISK = Auli Hakulinen, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, Irja Alho. 2004. *Ison suomen kieliopin verkkoversio* [The comprehensive grammar of Finnish, online version]. http://scripta.kotus.fi/visk/etusivu.php (accessed June 2014).
Zaikov, Pekka. 1992. *Karjalan kielen kielioppie* [A Grammar of Karelian]. Petroskoi: Izdateľstvo Petrozavodskogo gosudarstvennogo universiteta.

Kristina Kotcheva and Michael Rießler

Clausal complementation in Kildin, Skolt and North Saami

1 Introduction

The present chapter focuses on complementizer constructions with complement clauses and morphosyntactic devices identifying these constructions in three Saamic languages. Like in many other chapters of this volume, we focus on *canonical complementizers*, defined here as a subclass of complementizers marking finite complements, and we look specifically at the equivalents of English *that* and *if*. Complementizers marking non-finite complements as well as non-overt complement marking remain outside of the scope of the present paper and will not be described systematically.

The languages under investigation are Kildin Saami, Skolt Saami and North Saami. The structure of this chapter is as follows: Section 1.1 offers an overview on the investigated languages. Our choice for specifically these three languages is not arbitrary, but driven by considerations concerning available data and grammatical descriptions. These considerations will be discussed briefly in Section 1.2. Sections 2 and 3 represent the main discussion in our chapter. There we present the inventory of canonical complementation constructions in all three languages under investigation (Section 2). Based on this description, we discuss identified patterns of complementizer combination (Section 3.1) and omission (Section 3.2). Furthermore, we describe the multifunctionality (Section 3.3) and semantic types (Section 3.4) of complementizers. Section 4 consists of a short historical-comparative description of etymological sources for the formatives in all three languages discussed here. Although historical linguistics is outside the main scope of our investigation, these data might add interesting facts to this comparative volume on the syntax and semantics of complementation. Section 5 summarizes our findings.

Our data supports an analysis of complementizers in Kildin, Skolt and North Saami in terms of epistemic contrast. In Skolt and North Saami, we have on the one hand an epistemically neutral marker which does not pose any semantic restrictions upon the truth value of the proposition in the clause it introduces. On the other hand, we have an emerging complementizer used in cases where the truth of the proposition in its clause is uncertain. In Kildin Saami, the complementizers are better described as indicating certainty vs. uncertainty.

1.1 Saamic languages

The Saamic languages belong to the Finno-Ugric branch of the Uralic language family and form a dialect continuum across an area extending over northern Fennoscandia from central Scandinavia to the Kola Peninsula; see the map in Figure 1 and the language tree in Figure 2.

Typologically, Saamic languages are unlike the Indoeuropean languages spoken in the area by being almost exclusively suffixing languages and predominantly exhibiting postpositions. Nominal phrases are prototypically head-final, except for relative constructions which are postponed as a rule. Clause-level constituent ordering in Saamic is described as being free from formal restrictions, i.e. not determined syntactically, but pragmatically (Sammallahti 1998: 95; cf. also the discussion in Wilbur 2014: 226–230). However, SVO occurs more often in our corpus and might therefore be seen as the basic pattern for Kildin, Skolt and North Saami. In polar questions, the focused constituent (typically the verb) is fronted (Kildin) and additionally equipped with an enclitic question marker (North and Skolt Saami).

Another syntactic feature common to all Saamic languages is that negation is expressed by means of an inflected negation auxiliary followed by the main verb in a special connegative form (henceforth glossed as CN). Saamic case systems consist of between seven and nine cases and nouns additionally have possessive inflection (this feature is lost in Kildin Saami). In most Saamic languages, verbs and pronouns are inflected for dual (in Skolt dual is preserved only in pronominal inflection, in Kildin the dual is lost completely), in addition to singular and plural number values. Additionally, all Saamic languages are known for their complex non-linear morphology realized as so-called ablaut and consonant gradation (i.e. phonological alternations in a stem's vowels and consonants triggered by the morphological environment).

Due to the influence of Norwegian, Swedish, Finnish and Russian cultures and languages, all ethnic Saami are fluent in their respective contact languages, while younger generations are typically monolingual in these. As a result, all Saamic languages are today either extinct, moribund or endangered, but exhibit different levels of endangerment due to diverse language-sociological conditions and official support in the respective countries where these languages are spoken.

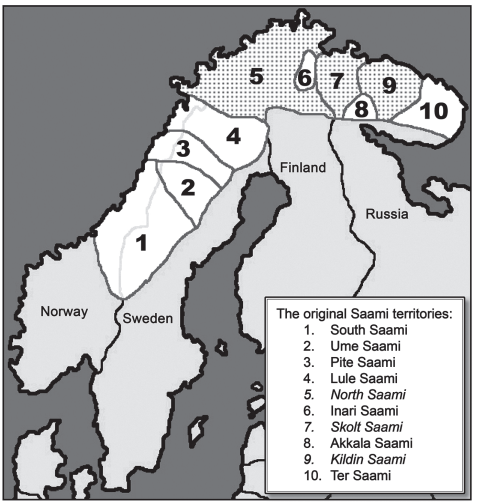


Figure 1: The Saamic languages and their traditional settlement areas; the geographic areas of the three investigated languages are shadowed

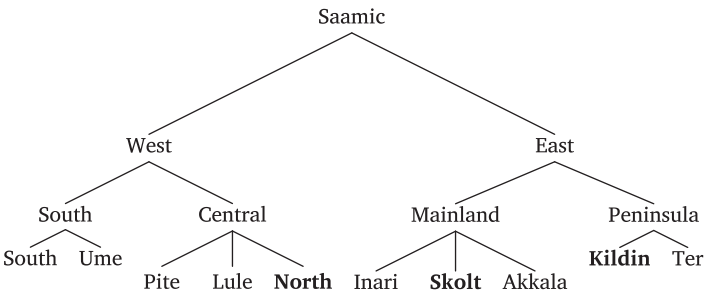


Figure 2: Saamic language tree according to Sammallahti (1998: 6–34); the three investigated languages are marked in bold face

1.1.1 Kildin Saami

Kildin Saami is not only the most endangered, but also the least documented and described among the three languages investigated here. Kildin Saami belongs to the peninsular subbranch of East-Saamic and is actively spoken by perhaps only 100 native speakers. Kildin Saami used to be spoken in the central inland parts and the central coastal parts of the Kola Peninsula, but the original Kildin dialect areas have fragmented, chiefly as the result of forced migration to larger towns.

As a result, most of the Kola Saami population lives in the town of Lovozero, which is nowadays usually regarded as the "Saami capital" of Russia.

A Cyrillic-based orthography for Kildin was developed during the 1980s and has been used in text books for elementary schools and in other teaching materials, including a few dictionaries. A considerable amount of literary texts have been published as well, although most of these are short or address only learners of Kildin Saami. The main contact language of contemporary Kildin Saami is Russian. Until recently there were also very close contacts with speakers of Karelian (cf. Blokland & Rießler 2011).

1.1.2 Skolt Saami

Skolt Saami is a member of the mainland subbranch of East-Saamic and has been spoken in Finland, Russia and Norway. The language used to be spoken predominantly on the western Kola Peninsula and the adjacent mainland in the borderland area of the three countries. However, most members of the original Skolt Saami villages in Russia resettled to Finland after the area was ceded to the Soviet Union in 1945. As a result, the language is now spoken in a relatively compact area by approximately 300 Saami in Finland. Skolt Saami in Finland speak either Suenjel or Petsamo dialect. In Russia, there are only about 30 Skolt Saami speakers left today, almost all of which speak the Notozero dialect. The traditional Skolt Saami dialect of Sør-Varanger in Norway is no longer spoken.

Resettlement to Finland is likely the explanation for the fact that Skolt Saami has been more extensively documented and described by Finnish linguists. It has also been successfully standardized in Finland and, due to support from the Finnish state, is probably somewhat less endangered than the closely related Kildin Saami language of Russia.

The main contact languages of contemporary Skolt Saami are Finnish (in Finland) and Russian (in Russia). Until recently there were also very close contacts with speakers of Karelian and Norwegian (cf. Blokland & Rießler 2011). Additionally, many Skolt Saami speakers in Finland speak North Saami. Skolt Saami has a standardized orthography based on Latin script and developed in Finland in the 1970s. This orthography could in principle cover all Skolt Saami dialects in the different countries (just as the standard North Saami orthography is valid in Finland, Norway and Sweden). However, the Skolt Saami written language has so far only been used in Finland.

1.1.3 North Saami

North Saami is the least endangered Saamic language. It belongs to the western subbranch and is spoken by about 17,000 speakers in Norway, Sweden and Finland (Sammallahti 1998: 1). North Saami is divided into several dialect groups and subgroups. The three main dialects are (from North to South) Sea North Saami, spoken along the northernmost Norwegian coast, Finnmark North Saami spoken in the inland of northernmost Norway and adjacent north-westernmost Finland, and Torne North Saami spoken in northernmost Sweden and adjacent areas in Norway and Finland (Sammallahti 1998: 9–20).

The main contact languages of contemporary North Saami in the respective countries are Finnish, Norwegian and Swedish. Additionally, many North Saami in Norway and Sweden speak Finnish. North Saami is relatively well described. Standard written North Saami is based on Latin script. The language is used in written media of all kinds today.[1]

1.2 Data sources

Our data are taken from existing descriptions and from available spoken and written corpora. The constructions identified in grammars and corpora were rechecked and additional data was elicited with native speakers. The data sources used for the present investigation are of diverse quantity and quality. This reflects the rather heterogeneous state of documentation and description of Saamic languages in general.

Comprehensive syntactic descriptions for Kildin Saami are virtually nonexistent. The sections on syntax in Kert's (1971) grammar present only relatively few examples. Therefore, most data on Kildin Saami is taken from a corpus of written and spoken texts collected and annotated by the Kola Saami Documentation Project (KSDP)[2] and its successor projects. Additional data was elicited for the present survey.

Skolt Saami is better described than Kildin Saami. A comprehensive grammatical description is Feist (2010). This descriptive grammar includes a discussion of complementation and other types of clausal subordination. A few examples are also included in the chapter on connectors in the school grammar by Moshnikoff, Moshnikoff and Koponen (2009). Additionally, we used examples

1 An overview on North Saami language sociology can be found in Aikio (2003). For Skolt and Kildin Saami, see Siegl and Rießler (2015).

2 http://dobes.mpi.nl/projects/sami (accessed 7 July 2014).

from recorded and annotated spoken texts available from KSDP or in elsewhere published text samples, among them the three transcribed spoken texts in the appendix of Feist's grammar. In addition, we have used elicited data.

For North Saami we used predominantly grammatical descriptions, the most important ones for us were Svonni (2009), Nickel (1994) and Nielsen (1979). Sammallahti (1998) includes a short section on subordination with a few examples. Nielsen's grammar, originally published in 1926–1929, is still considered the basic descriptive grammar of North Saami. Nickel's and Svonni's grammars are reference grammars describing the current standard language. In principle, there is also extensive digital corpus data available.[3] These corpora represent predominantly a standard written variety of the language and large parts of the relevant texts are North Saami translations of texts originally published in other languages. Thus the kind of North Saami data we have used differs considerably from Kildin and Skolt Saami data. Including North Saami in the present comparison nevertheless seems useful because North Saami is the most documented and best known Saamic language, and has been described in several comparative investigations previously.

When referring to data sources in the examples we apply the usual citation style for pieces of text one can find in printed sources. In square brackets we refer to recording sessions or other archived data corpora. Both kinds of corpus references are in italics, in contrast to normal references to scientific literature, such as reference grammars. Finally, for each example we also indicate whether the data originate from an elicitation session or whether the text in question is from a spoken or written register of the language in question.

Examples are represented in their respective contemporary orthographies, but note that Kildin Saami examples have been transliterated from Cyrillic to Latin script.

2 Inventory

In this section we present the inventory of canonical complementation constructions in each of the three languages under investigation and describe the syntactic environments in which they occur.

3 The most important corpus infrastructure for Saamic languages is provided by the Center for Sámi Language Technology (Giellatekno), see http://giellatekno.uit.no/.

2.1 Kildin Saami

Canonical complementizers in Kildin Saami are *šte* 'that' and *koal'e* or *jesl'e* 'if'. While the section on clausal subordination in Kert's descriptive grammar of Kildin Saami mentions only examples of *jesl'e* (Kert 1971: 286–296), the speakers from whom we obtained the major part of our spoken language corpus use *koal'e* almost exclusively. The attested variation between *koal'e* and *jesl'e* appears to be dialectal. There is also stylistic variation because *koal'e* occurs consistently in written texts and is also the main variant found in the three standard language dictionaries (Kert 1986: 152; Sammallahti & Chvorostuchina 1991: 26, 53; Afanas'eva et al. 1985).[4] The sociolinguistic determinants of the variation between *koal'e* and *jesl'e* are not the focus of the present investigation. In the following we consider both forms as free variants and use *koal'e* – occurring predominantly in our own data as well as in written standard language – as the main form.

Both *šte* and *koal'e* mark (direct) object complements, cf. *šte* in (1a), and *koal'e* in (1b). They introduce subject complement clauses as well, as shown in (2) where the subordinate clauses are complements of the adjective *šīg* 'good'.

(1) Direct object complement

a. *[...]* *tiedxell* *sōn* [***šte*** *vur'tse* *bidt*]
let.know her/him **šte** wait is.necessary
'(He) let him know that he needs to wait.'
(spoken *[KIL040300NEA_pearstory]*)

b. *sōnn* *tīd'e* [***koal'e*** *sōnn* *jo* *ujjtmal* *lī*]
s/he knew **koal'e** s/he already gone is
'S/he$_i$ knows if/whether s/he$_j$ has already gone.'
(elicited *[sjd201401NEA_compl]*)

(2) Subject complement

a. *adt'* *šīg* *lī* [***šte*** *ajkalt* *al'kev* *parnet'* *mat'xe*]
now good is **šte** early they.start children teach
'Today it is good that they start to teach the children early.'
(spoken *[KIL061201Katja]*)

b. *Šīg,* [***koal'e*** *puadt*].
good **koal'e** s/he.comes
'(It is) good if s/he comes.'
(written *[Afanas'eva et al. 1985: 410]*)

4 Kert (1986) lists *jesl'e* as a variant of *koal'*(*e*) in the Russian-Saami part of his dictionary, but neither form occurs in the Saami-Russian part. Also Sammallahti and Chvorostuchina (1991) provide *jesl'e* as a variant, but mark it as a Russian nonce-borrowing. Afanas'eva et al. (1985) do not include *koal'e* as a lemma, but there are several instances of the word among the example sentences for other lemmas; *jesl'e* is not used at all in this dictionary.

Beyond their use as canonical complementizers, *šte* and *koal'e* can also occur with adverbial clauses. *Koal'e* is the prototypical subjunctor marking conditional adverbial clauses in Kildin Saami; cf. example (3). In purpositive adverbial clauses, *šte* is frequently used along with the more common *štobe* 'so that'; cf. the examples in (4).

(3) *koal'e* as an adverbial subjunctor
koal'e *tijj* *ann'tbedt'e* *minn'e* *ājk* *munn* *jurr'tla* *tenn* *bajas*
koal'e you give me time I think that about
'If you give me some time, I'll think about it.'
(elicited *[sjd201401NEAcompl.012]*)

(4) *štobe* and *šte* as adverbial subjunctors
a. *milknes't* *pijj* ***štobe*** *sīnet'* *ell'a* *čūzxe*
slowly s/he.puts **štobe** them not.is hit.CN
'Slowly he puts (them) down in order not to smash (them).'
(spoken *[KIL040300NEA_pearstory]*)
b. *[...]* *milknes't* ***šte*** *ejj* *miejte* *sīnet'*
slowly **šte** not mash.CN them
'(He puts the pears down) slowly so that they don't get mashed.'
(spoken *[KIL040300NEA_pearstory]*)

Last but not least, predicates can also take complement clauses which are introduced by question words (like the pronouns 'what', 'who', the adverbs 'where', 'to where', 'from were', the quantifier 'how many', etc.); see examples (5a) and (5b) with the pronoun *mī* 'what' (accusative *mēnn*), (5c) with the pronoun *k'ē* 'who', and (5d) with the adverb *kas't* 'where'.

(5) a. Subject complement
čofta *tiġvenn* *lī* [***mēnn*** *kɨrj* *tōnn* *lōgak*]
very interesting is **what:ACC.SG** book.ACC.SG you read
'It's very interesting which book you are reading.'
(elicited *[sjd20140120NEA]*)
b. Direct object complement
munn *ujna* [***mēnn*** *kɨrj* *tōnn* *lōgak*]
I see **what:ACC.SG** book.ACC.SG you read
'I see which book you are reading.'
(elicited *[sjd20140120NEA]*)
c. Direct object complement
munn *ujna* [***k'ē*** *kɨrj* *lōgk*]
I see **who.NOM.SG** book.ACC.SG reads
'I see who reads a book.'
(elicited *[sjd20140120NEA]*)

d. Direct object complement
munn *ujna* [***kas't*** *tõnn* *lõgak* *kɨrj*]
I see **where** you read book.ACC.SG
'I see where you read the book.'
(elicited *[sjd20140120NEA]*)

Question words in these constructions are part of the subordinate clause and are syntactically governed according to their role within this clause. Consequently, the question words are not what actually identifies the complement as such. Therefore we are not analyzing them as canonical complementizers.

2.2 Skolt Saami

In Skolt Saami, *što* is mainly used as a *that*-type complementizer, but can even be interpreted as an *if*-type complementizer in some constructions. A variant of *što* is the recent Finnish loanword *että*. According to Feist (2010: 335), *što* and *että* can be used interchangeably, without any difference in the semantics of the relation they convey; see examples (6) and (7). However, this is true only for the Skolt Saami varieties spoken in Finland today. We could not find one single instance of *että* in the Skolt Saami texts recorded with speakers from the Russian side of the border. Interestingly enough, the Finish borrowing *että* never occurs in the Skolt Saami teaching materials created in Finland either. In the Skolt Saami reading book compiled from transcribed audio recordings of fairy tales and stories by Koponen, Moshnikoff and Moshnikoff (2010), all occurrences of *että* (as well as several other code-switches and nonce-borrowings from Finnish) are printed in italics in order to tag them as dispreferred Skolt Saami forms.

Presumably, *što* is considered original (and perhaps more "pure" Skolt Saami) and thus preferred by teachers and other authors of prescriptive teaching materials. The variation between the forms is thus not completely free, but determined dialectally and – in the two dialects spoken in Finland – also stylistically, with *što* being more characteristic of spoken Skolt Saami. Examples (6b) and (7b) provide occurrences in the standard language (in prefaces of teaching books) while examples (6a) and (7a) originate from spoken discourse (i.e. transcribed audio recordings of oral history).

(6) Subject complement
a. *saagg* *le´jje* [***että*** *puä´tte*]
news they.were **että** they.came]
'The news was that they came.'
(spoken *[Koponen, Moshnikoff and Moshnikoff 2010: 54]*)

b. *Nu´bb täävtõs lij,* [***što*** *rää´ves oummu še vuäitče*
second goal is **što** adult person also help
mätt'tõõttâd tän ǩee´rj mie´ldd ǩee´rjtum sää´mǩiõl].
teach this book with written Saami:language]
'The second goal is to familiarize Skolt Saami adults with their native language in written form.'
(written *[Moshnikoff, Moshnikoff and Koponen 2009: 7]*)

(7) Direct object complement
a. *Sverloov Jääkk jordd* [***että*** *mon sǩõlddõõđam*]
Sverloff's Jack thinks **että** I joke
'Jack Sverloff thinks that I'm joking.'
(spoken *[cit. Feist 2010: 335]*)
b. *Tuäivvap,* [***što*** *tän ǩee´rjest lij hää´sǩesvuõtt [...]*]
we.hope **što** this from.book is pleasure
'We hope that you find pleasure from this book.'
(written *[Koponen, Moshnikoff and Moshnikoff 2010: 6]*)

In Skolt Saami, the interchangeable *if*-type complementizers *jos*[5] and *je´sli* are attested as well. Both *što*/*että* and *jos*/*je´sli* can be used with subject and object clausal complements; cf. the minimal pairs in (8a)–(8c) and (9a)–(9c).

(8) Subject complement
a. *šiõgg lij* [***što*** *son puätt*]
good is **što** s/he comes
'It is good that s/he comes.'
(elicited *[sms20130614ZMN, sms20140616EL]*)
b. *šiõgg lij* [***je´sli*** *son puätt*]
good is **je´sli** s/he comes
'It is good if s/he comes.'
(elicited *[sms20130614ZMN]*)
c. *šiõgg lij* [***jos*** *son puätt*]
good is **jos** s/he comes
'It is good if s/he comes.'
(elicited *[sms20140616EL]*)

(9) Object complement
a. *son teâtt* [***što*** *autt lij puättam*]
s/he knows **što** car is come
'S/he knows that the car has arrived.'
(elicited *[sms20130614ZMN, sms20140616EL]*)
b. *son teâtt* [***je´sli*** *autt lij puättam*]
s/he knows **je´sli** car is come
S/he knows if the car has arrived.'
(elicited *[sms20130614ZMN]*)

5 The subordinator *jos* has a variant, *jõs*, which is less close to the Standard Finnish pronunciation (Feist 2010: 346), but is rarely attested in our data.

c. *son* *teâtt* [***jos*** *autt* *lij* *puättam*]
s/he knows **jos** car is come
S/he knows if the car has arrived.'
(elicited *[sms20140616EL]*)

The markers *jos* and *je´sli* are not cognates, but independent recent borrowings from Finnish and Russian, respectively: *jos* is used exclusively by speakers living in Finland, and *je´sli* only in Russia. The phonological resemblance between them is coincidental (see also Section 4).

All examples of the use of *jos* and *je´sli* as complementizers originate from elicitation sessions. In our spoken language corpus, we could only attest instances of *jos* and *je´sli* in conditional clauses. The same is true for our written language corpus and for the data used in the grammatical description and Skolt Saami texts provided by Feist (2010).

(10) *jiõm* *â´te* *mon* *ni* *kõõjjče* [***jos*** *mon* *teâđčem*]
I.not then I not would.ask.CN **jos** I would.know]
'I wouldn't even ask if I knew.'
(spoken *[cit. Feist 2010: 269]*)

Interestingly, Skolt Saami allows for *što* occurring in *if*-type complements, as in example (11).

(11) *ri´mjj-kää´lles* *vuõ´lji* *tõn* *tollsââ´jes* *ǩičč̆âd* [***što*** *lij=**a*** *aiham* *piâssâm*
fox-gaffer left that fireplace see **što** is=**PQ** bear escaped
le´be *puâllam* *avi* *mâi´d*]
or burned or what
'Mr. Fox left to see that fire place, (to see) whether Bear has escaped or burnt or whatever.'
(spoken *[Sammallahti 2012: 28]*)

That- and *if*-uses of the complementizer *što* are disambiguated by the syntactic structure of the complement clause. In (11), the complement clause displays a linear word order characteristic of polar questions, with the finite verb placed before the subject and marked with the question clitic *=a*. In the example (9a) above, in which *što* has a *that*-reading, the subject precedes the finite verb. This is the canonical word order of declarative clauses.

Just as in Kildin Saami, finite complement clauses can be introduced by question words. Unlike in Kildin Saami, such complement clauses can additionally be marked with the *that*-type complementizer *što*; see (12a) and (12b).

(12) a. Complement clause introduced by question adverb 'how'
Čuä´jet *mu´nne,* [*[**što**]* ***mä´htt*** *rabblin* *kåå´đet*].
one.shows me (**što**) **how** with.comb one.weaves
'It is shown to me how to weave with the (weaving) comb.'
(written *[Koponen, Moshnikoff and Moshnikoff 2010: 166]*)

b. Complement clause introduced by question adverb 'when'
Škoolnee´ǩǩ *jie* *tie´đ,* [***[što]*** ***kuä´ss*** *tä´lvvluõvâsvuõtt* *älgg*].
pupils they.not know.CN (**što**) **when** winter.holiday starts
'The pupils don't know when the winter holiday starts.'
(written *[Koponen, Moshnikoff and Moshnikoff 2010: 166]*)

The insertion of the complementizer is optional, as indicated by the parentheses in the examples. We shall deal with this kind of combination in more detail in Section 3.1.

2.3 North Saami

Canonical complementizers in North Saami are *ahte* 'that' and *jus*[6] 'if'. Typically, they serve as markers of subject complements, as in (13), or of (direct) object complements, as in (14).

(13) Subject complement
a. *Buorre* *lei,* [***ahte*** *nuorabut* *vulge* *dasto* *dán* *geasi*
good was **ahte** young.ones drew after that summer
ordnet *gilvvuid*].
organize competitions
'It was good that the youngsters organized a competition that summer.'
(written *[YLE.fi, 5.7.2011[7]]*)
b. *Buorre* *livččii* [***jus*** *stáhta* *livččii* *boraspiriid* *geahpeadeaddji*].
good would.be **jus** state would.be predators reduce
'It would be good if the state reduced (the number) of predator animals.'
(written *[YLE.fi, 2.2.2011[8]]*)

(14) Object complement
a. *Mun* *dieđán,* [***ahte*** *Niillas* *boahtá*].
I know **ahte** Nils comes
'I know that Nils is coming.'
(written *[Sammallahti 1998: 105]*)
b. *Mon* *in* *dieđe* [***jus*** *mon* *gearggan* *boahtit*].
I I.not know.CN **jus** I manage come
'I don't know if I'll manage to come.'
(written *[Svonni 2009: 92]*)

6 Two functionally equivalent spelling variants are *jos* and *juos*.
7 http: //yle.fi/uutiset/vuovdaguoikka_guolastangilvvuin_stuorra_vuoitoguolli/6625939 (accessed 22 October 2012)
8 http: //yle.fi/uutiset/spiriid_geahpedeapmi_heivesii_stahta_bargun/6625193 (accessed 22 October 2012)

Besides its use in complementation, *jus* introduces conditional clauses as well. Hypothetical conditionals as in (15a) are marked for conditional mood, simple conditionals as in (15b) are in indicative mood.

(15) a. Hypothetical conditional
*Mon jugašin gáfe **jus** gearggašin.*
I would.drink coffee **jus** I.would.manage
'I would drink coffee if I managed.'
(written *[Svonni 2009: 56]*)
b. Simple conditional
***Jus** ihttin lea buolaš, de eat vuolgge čuoigat.*
jus tomorrow is frost than we.not leave.CN ski
'If it is cold tomorrow we will not go skiing.'
(written *[Neahttadigisánit[9]]*)

As in Kildin and Skolt Saami, predicates can also take complement clauses which are introduced by question words, as illustrated by (16). We do not consider the question words to be canonical complementizers. Like in Skolt Saami (but unlike in Kildin Saami), such complement clauses are often marked optionally by adding the complementizer *ahte*, thus yielding *ahte maid* 'that which' in (17).

(16) *Son jearrá [**guhte** lea boahtán].*
s/he asks **who** is come
'S/he asks who has come.'
(written *[Svonni 2009: 88]*)

(17) *Muhto mii dušše geahčastalaime vuorrasii, [**ahte**] **maid** dal vuoras dahká].*
but we only looked at.old.man (**ahte**) **what:ACC.SG** now old.man does
'But we just looked at the old man: what is he going to do now?'
(written *[Nickel 1994: 196]*)

We shall deal with this kind of combination in more detail in Section 3.1.

3 Syntax and Semantics

Building on the inventory of canonical complementizers presented in the previous section, the present section describes possible complementizer combinations and omissions as well as the degree of multifunctionality of the identified complementizers. Furthermore, we present a typology of canonical complementation

9 http: //sánit.oahpa.no (accessed 22 June 2014)

constructions by using the equivalents of Noonan's (2007) prototypical classes of complement-taking predicates.

3.1 Combination

In all three Saamic languages, complement clauses of predicates of negative propositional attitude or of utterance predicates like 'ask' or 'tell' introduced by the *if*-type complementizer are equivalent to embedded polar questions.

(18) *if*-type complementation in Kildin Saami
a. M*un* *emm* *t'ēd'*, [***koal'e*** *Evvan* *lī* *puadtma*].
I I.not know.CN **koal'e** Ivan is come
'I don't know if/whether Ivan has come.'
b. *Mun* *emm* *t'ēd'*, [*lī* *Evvan* *puadtma*].
I I.not know.CN **Ø** is Ivan come
'I don't know if/whether Ivan has come' (I don't know, has Ivan come).

(19) *that*-type complementation in Kildin Saami
Mun *emm* *t'ēd'*, [***što*** *Evvan* *lī* *puadtma*].
I I.not know.CN **što** Ivan is come
'I don't know that Ivan has come.'

(20) *if*-type complementation in Skolt Saami
a. *Mon* *jiõm* *tie´đ*, [***jos*** *autt* *lij* *puättam*].
I I.not know.CN **jos** car is come
'I don't know if/whether the car has come.'
b. *Mon* *jiõm* *tie´đ*, [*lij=**a*** *autt* *puättam*].
I I.not know.CN is=**PQ** car come
'I don't know if/whether the car has come' (I don't know, has the car come).

(21) *that*-type complementation in Skolt Saami
Mon *jiõm* *tie´đ*, [***što*** *autt* *lij* *puättam*].
I I.not know.CN **što** car is come
'I don't know that the car has come.'

(22) *if*-type complementation in North Saami
a. *Ante* *ij* *dieđe*, [***jus*** *Máhtte* *áigu* *njuovvat* *sávzza*]
Andrew s/he.not know.CN **jus** Matthew will butcher sheep
'Andrew doesn't know if/whether Matthew is going to butcher a sheep.'
b. *Ante* *ij* *dieđe*, [*áigu=**go*** *Máhtte* *njuovvat* *sávzza*]
Andrew s/he.not know.CN will=**PQ** Matthew butcher sheep
'Andrew doesn't know if/whether Matthew is going to butcher a sheep' (Andrew doesn't know, is Matthew going to butcher a sheep).

(23) *that*-type complementation in North Saami
Ante *ij* *dieđe,* [***ahte*** *Máhtte* *áigu* *njuovvat* *sávzza*]
Andrew s/he.not know.CN **ahte** Matthew will butcher sheep
'Andrew doesn't know that Matthew is going to butcher a sheep.'

This results in a minimal triplet if the two options for marking *if*-complements are contrasted to the regular *that*-type complementation in (19), (21), and (23). However, the tripartite contrast is valid only with regard to the construction; semantically it is a bipartite contrast between the *that*- and *if*-type complementation; cf. (18), (20), and (22).

Like in Kildin Saami (18b), the subordinated clauses in Skolt (20b) and North Saami (22b) can be analyzed as juxtaposed polar questions. Unlike in Kildin Saami, where polar questions are marked by intonation alone, polar questions in Skolt and North Saami are obligatorily marked by means of a question enclitic.[10] We do not analyze the relevant enclitics North Saami *=go* and Skolt Saami *=a* as complementizers but rather consider them to be regular question markers. The complement construction is overtly unmarked because there is no additional formative identifying the complement as such.

In Skolt and North Saami, though not in Kildin Saami, the respective *that*-type complementizers *što/ahte* are frequently inserted before the embedded polar question, cf. (24) and (25).

(24) Skolt Saami
a. *Mon* *jiõm* *silttâd* *särnnad,* [*[**što**]* *puätt=**a*** *son* *avi* *ij*].
I I.not can.CN tell (**što**) comes=**PQ** s/he or s/he.not
'I cannot tell whether or not s/he will come.'
(written *[Koponen, Moshnikoff and Moshnikoff 2010: 166]*)
b. *Mä´rjj* *kõõjji* *Jakkust,* [*[**što**]* *leäi=**a*** *son* *kåddam* *kue´l* *tõn*
Mary asked from.Jack (**što**) was=**PQ** s/he catched fish this
jääu´rest].
from.lake
'Mary asked Jack if/whether he has caught fish in this lake.'
(written *[Koponen, Moshnikoff and Moshnikoff 2010: 166]*)

(25) North Saami
a. *Aigomarvalit* [*[**ahte**]* *eat=**go*** *mii* *ihtin* *vuolgi*].
question (**ahte**) we.not=**PQ** we tomorrow leave.CN
'I will ask if/whether we can't leave tomorrow.' (written *[Nielsen 1979: 404]*)
b. *Dat* *jerre,* [*[**ahte**]* *máhtán=**go*** *mun* *dan* *luođi*].
it asked (**ahte**) can=**PQ** I this song
'They asked if/whether I can (sing) this song (joik).'
(written *[Nickel 1994: 437]*)

10 Word order is also relevant in all three languages because the focused constituent is always fronted.

Our Kildin Saami corpus does not provide similar examples. According to our native speaker consultants, inserting *šte* in such constructions is disallowed in Kildin Saami.

North and Skolt Saami, but not Kildin Saami, allow also for *ahte/što* in front of complement clauses introduced by wh-question words. Skolt Saami examples were given above in (12), repeated below as (26); a North Saami example is (27).

(26) (=12) Skolt Saami

a. Complement clause introduced by question adverb 'how'
Čuä´jet *mu´nne,* [***[što]*** ***mä´htt*** *rabblin* *kåå´đet*].
one.shows me (**što**) **how** with.comb one.weaves
'It is shown to me how to weave with the (weaving) comb.'
(written *[Koponen, Moshnikoff and Moshnikoff 2010: 166]*)

b. Complement clause introduced by question adverb 'when'
Škooulnee´ǩǩ *jie* *tie´đ,* [***[što]*** ***kuä´ss*** *tä´lvvluõvâsvuõtt* *älgg*].
pupils they.not know.CN (**što**) **when** winter.holiday starts
'The pupils don't know when the winter holiday starts.'
(written *[Koponen, Moshnikoff and Moshnikoff 2010: 166]*)

(27) North Saami

a. *Diehtá* [***[atte]*** ***mii*** *dalle* *lea*].
know (**ahte**) **what:NOM.SG** there is
'(H/She) knows what that over there is.'
(written *[Nielsen 1979: 404]*)

b. *Áidna* *gážaldat* *dalle* *lea* [***[ahte]*** ***maid*** *sii* *galget* *geat*
only.one question then is (**ahte**) **what:ACC.SG** they should who:NOM.PL
guđđet *ealáhusa*].
abandon economy
'The only question then is what those who abandon the economy should do.'
(written *[NRK.no*[11]*]*)

Since Skolt and North Saami *što/ahte* cannot be dropped in constructions such as those in (21) and (23), example (27) illustrates insertion of *što/ahte* in front of question-like complement clauses rather than complementizer omission.

In Section 3.3 we discuss the use of *šte/što/ahte* in constructions other than canonical complementation.

3.2 Omission

According to the grammatical descriptions, omission of the *that*-type complementizer is a possible variant in North Saami; see example (28) from Nickel (1994).

11 http: //www.nrk.no/kanal/nrk_sapmi/samegillii/1.10899516 (accessed 20 April 2014)

(28) North Saami
Máhtte *muitalii* [***[ahte]*** *áddjá* *boahtá*].
Matthew told (**ahte**) grandfather comes
'Matthew said (that) grandfather is coming.'
(written *[Nickel 1994: 439]*)

In Skolt and Kildin Saami, this kind of variation is regularly found only in cases of direct quotation, although at least in spoken varieties of Skolt Saami there is a noticeable tendency to mark a direct citation overtly with the complementizer.

Strictly speaking, direct citation (and reported speech) belongs to canonical complementation since the subordinate clause (typically including a finite predicate) serves as the direct object of an utterance verb. In (29a) the predicate in the clause introduced by *što* – the verb *vääldam* 'I take' – is marked for first person singular, thus clearly indicating that we have quoted rather than reported speech. In the latter case, the predicate would be marked for third person. In example (29b), Skolt Saami *što* is used in this function twice, introducing two direct quotations nested into one another;[12] see also the double quotes in the translation.

(29) Skolt Saami
a. *ooumaž [...]* *ceälkk* [***što*** *no* *peä´l* *vääldam,* *peä´l* *jällmõõžž* *vääldam*]
person says **što** well half I.take half life I.take
'The man [...] says: "Well, I'll take a half, the half life I'll take."'
(spoken *[Koponen, Moshnikoff and Moshnikoff 2010: 13]*)
b. *mon* [***što*** *no* *ton* *cie´lǩǩiǩ* [***što*** *kuä´mmertiudd* *ooumže*]]
I **što** well you said **što** handful to.person
'I (said): "Well, you said: 'One handful per person.'"
(spoken *[cit. Feist 2010: 233]*)

The use of the *that*-type complementizer for marking direct quotation is most pervasively attested in Skolt Saami (cf. Feist 2010: 337–338), but the marker is also regularly used in this function in Kildin and North Saami (cf. Kert 1971: 295–296 for Kildin Saami and Nielsen 1979: 405–406 for North Saami). Note that reported speech can also be subordinated by pure juxtaposition in all three languages. This is usually the case in written varieties since direct quotations are marked typographically (with the embedded clause in quotation marks).

12 Note also the non-expressed (but implied) utterance verb 'said' in the matrix clause in example (29b).

3.3 Multifunctionality

Kildin Saami *šte*, Skolt Saami *što*/*että* and North Saami *ahte* introduce various subordinate constructions beyond canonical complementation and can therefore be described as multifunctional subjunctors. However, the extent to which the complementizers are multifunctional differs between the individual languages. Skolt Saami *što*/*että* displays the widest range of uses. It allows for both *that*- and *if*-type interpretations (see example 11 in Section 2.2) and introduces direct quotations, adverbial clauses and other subordinated constructions. The *that*-type complementizer *ahte* in North Saami occurs in similar constructions, but less regularly than its Skolt Saami counterpart. Kildin Saami *šte* is also multifunctional, but to an even lesser degree. In the following, we shall present examples of the multifunctionality of complementizers.

3.3.1 Adverbial clause subordination

In all three languages, the *if*-type complementizer introduces adverbial conditional clauses.

(30) a. (=3) Kildin Saami

koal'e	*tijj*	*ann'tbedt'e*	*mɨnn'e*	*ājk*	*munn*	*jurr'tla*	*tenn*	*bajas*
koal'e	you	give	me	time	I	think	that	about

'If you give me some time, I'll think about it.'
(elicited *[sjd201401NEAcompl.012]*)

b. (=10) Skolt Saami

jiõm	*â´te*	*mon*	*ni*	*kõõjjče*	***jos***	*mon*	*teâđčem*
I.not	then	I	not	not.would.ask.CN	**jos**	I	would.know

'I wouldn't even ask if I knew.'
(spoken *[cit. Feist 2010: 269]*)

c. North Saami

Leat	*álbmotoaju*	*miellahttun*	***jus***	*orut*	*Norggas.*
you.are	social.security	member	**jus**	you.live	in.Norway

'You are member of the social security system if you live in Norway.'
(written *[Nav.no*[13]*]*)

Kildin Saami *šte*, Skolt Saami *što*/*että* and North Saami *ahte* are also used in adverbial subordination; cf. the Kildin Saami examples in (31), where an adverbial clause of reason is subordinated by means of *šte*. This use is characteristic for both spoken and written styles.

13 Cited from the Giellatekno Corpus, https://victorio.uit.no/freecorpus/orig/sme/facta/nav.no/medlemskap_i_folketrygden.html (accessed 9 August 2013)

(31) Kildin Saami[14]

a. *sōnn* *jurtel'* *šte* *sām'* *pudt'en'* ***šte*** *sōn* *narodt* *ell'aj* *il'm'en'*
s/he$_i$ thought šte Saami came **šte** her/his$_j$ people not.is living
'S/he thought that the Saami came because his/her people are dead.'
(spoken *[Ėndjukovskij 1937: 163]*)

b. *Pāss'pe* *kānnca,* ***šte*** *viek'eht'.*
thanks to.friend **šte** s/he.helped
'Thanks to the friend because s/he helped.'
(written *[Afanas'eva et al. 1985: 250]*)

In all three languages the *that*-type complementizer can introduce a subordinated adverbial clause of purpose as well; cf. examples (32a)–(32c).

(32) a. (=4b) Kildin Saami
[...] milknes't ***šte*** *ejj* *miejte* *sīnet'*
slowly **šte** s/he.not mash.CN them
'(He puts the pears down) slowly so that they don't get mashed.'
(spoken *[KIL040300NEA_pearstory]*)

b. Skolt Saami
Mee´st *kâ´l* *tõk* *jeä* *leämmaš,* ***što*** *igõl*
at.us arguably those they.not were.CN **što** not.need
säämas *särnnad.*
in.Saami speak
'We didn't have any of those, so that we didn't need to speak Saami.'
(spoken *[cit. Feist 2010: 443]*)

c. North Saami
Niilas *vulggii,* ***ahte*** *mun* *bessen* *nohkat.*
Nils left **ahte** I was.free sleep
'Nils left so that I could go to sleep.'
(written *[Sammallahti and Nickel 2006: 3]*)

3.3.2 Non-verbal predicates

The *that*-type complementizers are also frequently used in constructions in which the subordinate clause is dependent on an adpositional or adverbial phrase inside the matrix clause. These constructions resemble relative clause constructions because the subject of the dependent clause is co-referent with the syntactic head in the matrix clause: an adverb in (33a) and indefinite pronouns in (33b) and (33c).

14 Example (31a) shows two occurrences of *šte*, the first illustrating its complementizer, the second its adverbializer use.

(33) a. Kildin Saami
Vijt'e ***nɨdt', šte*** ...
turned.out **so** **šte** ...
'(It) turned out such that [...]'
(written *[Afanas'eva et al. 1985: 218]*)

b. Skolt Saami
mutta *puäʒʒ-han* *lij* ***nåkam*** ***što*** *tõt* *šâdd* *sõrgg*
but reindeer-also is **such** **što** this grows immediately
'But a reindeer is also such that it is growing immediately.'
(spoken *[Koponen, Moshnikoff and Moshnikoff 2010: 34]*)

c. North Saami
Dat *visti* *lea* ***dakkár,*** ***ahte*** *das* *lea* *feaskkir.*
this house is **such** **ahte** in.it is mud.room
'This house is such that is has a mud room.'
(written *[Nielsen 1979: 194]*)

That-type complementizers used in this kind of construction tend to be reanalyzed as subordination markers, as evidence from several languages demonstrates; cf. German *so dass* > *sodass* (so+that) 'so that' or Russian *potomu, čto* > *potomu čto* > *potomučto* (for+this+that) 'because'. In fact, Kildin Saami *nɨdt'*, *šte* 'so that' has already been reanalyzed as confirmed by its frequent occurrence in texts and by the native speaker intuition of our consultants. The examples presented above are nevertheless valid for our argumentation that Kildin, Skolt, and North Saami *that*-type complementizers are multifunctional.

3.4 Semantics

Similar to Germanic, the Saamic languages under scrutiny distinguish between formatives introducing finite clausal complements, roughly equivalent to *that*- and *if*-type complementizers in English.

Generally, complement-taking predicates can be divided into three groups with respect to the complementizers used to introduce a finite complement clause: a) predicates allowing only for *that*-type complementizers, e.g. German *glauben* 'to believe'; b) predicates allowing only for *if*-type complementizers, usually interrogative utterance predicates, or verbs like German *überlegen* 'to reflect', and c) predicates allowing for both *that*- and *if*-type complementizers, e.g. German *wissen* 'to know' (Bäuerle & Zimmermann 1991; cf. also Nordström & Boye, this volume). At first glance, the same seems to be true for Kildin, Skolt and North Saami.

With predicates expressing positive propositional attitudes such as 'believe', 'think', 'suppose', or 'assume', only *that*-type complementizers were attested. In affirmative matrix clauses, such as those in examples (34a)–(34c), the proposi-

tion in the complement clause is asserted to be true from the perspective of the subject referent. In negated matrix clauses, the proposition in the complement clause is asserted to be false from the perspective of the subject referent.

(34) a. Kildin Saami
sōnn *tennxanna* *jurr'tel* [***šte*** *sōnn* *vjāl* *puadt*]
s/he$_i$ immediately thought **šte** s/he$_j$ again comes
'S/he$_i$ immediately thought that s/he$_j$ will come again.'
(elicited *[sjd201401NEAcompl.011]*)

b. Skolt Saami
[...] *čuõškk* *jordd* [***što*** *pakk* *Ǩie´sspiõgg* *pååss*]
mosquito thinks **što** warm summer.wind blows]
'The mosquito thinks that a warm summer wind is blowing.'
(spoken *[cit. Feist 2010: 380]*)

c. North Saami
mii *doaivut* [***ahte*** *sii* *leat* *boahtán*]
we think **ahte** they are come]
'We think that they have arrived.'
(elicited *[sme20140101comp]*)

Predicates of knowledge and acquisition of knowledge, such as 'know', 'learn', 'discover' as well as perception predicates like 'see' or 'hear', allow complement clauses to be introduced by both *that*- and *if*-type complementizers in all three Saami languages.

(35) Kildin Saami
a. *sōnn* *ejj* *t'ēdtma* [***šte*** *sōnn* *jo* *ujjtma* *lī*]
s/he s/he.not know.CN **šte** s/he already gone is
'S/he doesn't know that s/he was already gone.' (elicited *[sjd201401NEAcompl.048]*)

b. *sōnn* *ejj* *t'ēdtma* [***koal'e*** *sōnn* *jo* *ujjtma* *lī*]
s/he s/he.not know.CN **koal'e** s/he already gone is
'S/he doesn't know if/whether s/he was already gone.'
(elicited *[sjd201401NEAcompl.048]*)

(36) Skolt Saami
a. *vuõi´ni,* [***što*** *ääldast* *leäi* *pâi* *võrrpääikaž* *pääccam*]
saw **što** from.female.reindeer was only small.bloodplace remained
'(S/he) saw that only a patch of blood had remained of the reindeer.'
(spoken *[cit. Feist 2010: 256]*)

b. *mon* *jiõm* *tie´đ,* [***jos*** *autt* *lij* *puättam*]
I I.not know.CN **jos** car is come
'I don't know if/whether the car has come.'
(elicited *[sms20140616EL]*)

(37) North Saami

a. *Mun oainnán,* [***ahte*** *áddjá* *lea* *boahtán*].
I see **ahte** grandfather is come
'I see that grandfather has come.'
(written *[Nickel 1994: 194]*)

b. *Mun oainnán,* [***jus*** *son* *áigu* *njuovvat* *sávzza*].
I see **jus** s/he will butcher sheep
'I shall see if/whether s/he will butcher a sheep.'
(elicited *[sme20140101comp]*)

In examples (35a), (36a), and (37a) above, the proposition in the complement clause introduced by a *that*-type complementizer is asserted to be true from the perspective of the speaker uttering the complex sentence, irrespective of any negation in the matrix clause. On the other hand, the truth value of the propositions in clauses introduced by an *if*-type complementizer in examples (35b), (36b), and (37b) remains undetermined, both as part of an affirmative and a negated matrix clause. With the *if*-type complementizer, both the subject referent of the matrix clause and the speaker uttering the complex sentence are ignorant of the truth value of the complement clause's proposition. Thus the *if*-type complementizer can be analyzed as indicating a speaker or subject referent's uncertainty with regard to the truth value of the clause of which it is part.

Interrogative utterance predicates only allow complement clauses with a complementizer of the *if*-type; cf.examples (38a)–(38c).

(38) a. Kildin Saami
sõnn *kẽž'el'* [***koal'e*** *sõnn* *ujjn* *pūdze*]
s/he$_i$ asked **koal'e** s/he$_j$ sees reindeer
'S/he$_i$ asked if s/he$_j$ sees the reindeer.'
(elicited *[sjd20140120NEA]*)

b. Skolt Saami
son *kõõjji* [***jos*** *mon* *vuäinam* *puõccu*]
s/he asked **jos** I see reindeer
'S/he asked if I see a reindeer.'
(elicited *[sms20130614ZMN]*)

c. North Saami
[...] *sii* *leat* *jearran* [***jos*** *mun* *sáhtán* *oahpahit* *sámedánssu*]
they are asked **jos** I can teach Saami.dance
'They have asked if I can teach Saami dance.'
(written *[Ville.se*[15]*]*)

In complement clauses of interrogative utterance predicates, the truth value of the proposition in the dependent clause remains undetermined. Thus complement clauses introduced by the *if*-type complementizer share pivotal features

15 http: //villle.blogspot.de/2006_06_01_archive.html (accessed 22 June 2014)

with interrogative clauses, especially with polar interrogative clauses. In polar questions, the subject referent requires the confirmation or rejection of the proposition in question.

Skolt and North Saami, but not Kildin Saami, allow complement clauses introduced by an *if*-type complementizer to be substituted by a polar question, cf. Skolt Saami (24a) from Section 3.1 repeated here as (39).

(39) (=24) Skolt Saami

Mon	*jiõm*	*silttâd*	*särnnad,*	[*puätt=**a***	*son*	*avi*	*ij*].
I	I.not	can.CN	tell	comes=**PQ**	s/he	or	s/he.not

'I cannot tell whether or not s/he will come.'
(written *[Koponen, Moshnikoff and Moshnikoff 2010: 166]*)

There is therefore an attested semantic affinity between complement clauses introduced by an *if*-type complementizer and polar questions, as defined in (40).

(40) *if*-type complementizer + declarative clause = zero complementizer + polar interrogative clause

This seems to support an analysis of two different types of complementation marking in Saamic in terms of differences in illocutionary force in which the *if*-type complementizer conveys an interrogative illocutionary force in being an "interrogative" complementizer. The *that*-type complementizer in turn would be analyzed as a "declarative" complementizer and is therefore expected to be banned from use with non-declarative/non-assertive clauses.

However, the fact that in Skolt and North Saami, the respective *that*-type complementizers are compatible with question-formed complement clauses contradicts such an analysis; cf. (41).

(41) Skolt Saami

sij	*kõ´čče*	*su´st,*	[***što***	*ku´ǩǩen=**a***	*lij*	*sää´msijdd*]
they	asked	from.them	**što**	far.away=**PQ**	is	Saami.village

'They asked if the Skolt Saami village is far away.'
(spoken *[cit. Feist 2010: 337]*)

Thus, we can replace the definition in (40) with (42).

(42) *if*-type complementizer + declarative clause = zero or *that*-type complementizer + polar interrogative clause

This seems puzzling if we assume that at least in Skolt and North Saami the *that*-type complementizer is used to indicate that the proposition in its clause holds either from the perspective of the subject referent or from the perspective of the speaker. Moreover, as we have already seen, the *that*-type complementizer in

Skolt and North Saami occurs in front of a complement clause introduced by a wh-question word as well.

(43) (=12b) Skolt Saami
*Škoolnee´ǩǩ jie tie´đ, [**[što]** **kuä´ss** tä´lvvluõvâsvuõtt älgg].*
pupils they.not know.CN **(što)** **when** winter.holiday starts
'The pupils don't know when the winter holiday starts.'
(written *[Koponen, Moshnikoff and Moshnikoff 2010: 166]*)

It is even more interesting that the *that*-type complementizers in Skolt and North Saami can be combined with the respective *if*-type complementizers, yielding a complex complementizer; see the combination of *što* and *jos* in Skolt Saami (44).

(44) *Skolt Saami*
*jiõm â´tte mon ni kõõjjče [**što** **jos** mon teâđčem [...]]*
I.not then I not would.ask.CN **što** **jos** I would.know
'I wouldn't even ask if I knew [...]'
(written *[cit. Feist 2010: 348]*)

Such constructions with combined *that*- and *if*-type complementizers are not attested in Kildin Saami.

The *that*-type complementizer in Skolt and North Saami can thus be used with non-assertive clauses; use of the label "declarative" for describing its properties is inappropriate. Moreover, an 'or what'-tag can be added at least to Skolt Saami *što*, as shown in example (11), repeated below as (45). This has been claimed to be possible only with *if*-type (i.e. "interrogative" type) complementizers (cf. Nordström & Boye, this volume, for Germanic).

(45) (=11) Skolt Saami
*ri´mjj-kää´lles vuõ´lji tõn tollsââ´jes ǩičč̂âd [**što** lij=**a** aiham piâssâm*
fox-gaffer left that fireplace see **što** is=**PQ** bear escaped
*le´be puâllam **avi** **mâi´d**]*
or burned **or** **what**
'Mr. Fox left that fire place to see whether Bear has escaped or burned, or whatever.'
(spoken *[Sammallahti 2012: 28]*)

Our data supports an analysis of Skolt Saami *što* and North Saami *ahte* as neutral subordinators. We prefer the term *subjunctor* over *complementizer* in this case since *što/ahte* do not occur exclusively with complement clauses, but with adverbial clauses as well.

The data presented in our survey are consistent with the analysis of Saamic *if*-type complementizers as uncertainty markers. Especially the possibility of combining *što* with *jos* in Skolt Saami points in this direction. Here, the complementizer *što* marks the clause as a dependent one, while *jos* adds the element of uncertainty.

The Skolt *što* and North Saami *ahte* are in fact compatible with all semantic classes of complement-taking predicates, receiving the reading of a *that*- or an *if*-type complementizer depending on the properties of the complement clause. In our Kildin Saami data, on the other hand, there is no attestation of *šte* combined with either *koal'e* or question words or its insertion in embedded polar questions. A possible explanation for this can be found in the fact that the marker *šte* is a transparent borrowing from Russian *čto* (see Section 4 below). Kildin Saami speakers, who have all been bilingual in Russian at least since the extensive Russification in the post-war Soviet Union (cf. Rießler 2009) might avoid such a pleonastic use of *šte* because it is not possible in Russian (see Hansen, Letuchiy & Błaszczyk, this volume). The bulk of Skolt Saami speakers, on the other hand, have been separated from a Russian speaking environment since their resettling to Finland during World War II. Note that our Skolt Saami native speaker consultant from the Russian side has also rejected the combined constructions with *šte*.

As a result, our analysis of the Kildin Saami complementizers *šte* and *koal'e* necessarily differs from the analysis provided above for Skolt and North Saami, where a neutral subjunctor is also used in complementation if the meaning of uncertainty is indicated by other means (such as by differing word order from the one in declarative clauses, or by a question particle). In Kildin Saami, we instead have the dichotomy of the complementizer *šte* indicating certainty, and the complementizer *koal'e* indicating uncertainty.

4 Evolution

The ultimate origin and evolution of complementation marking in Saamic remains unclear. We do not know how *that*-type complementation was encoded in West-Saamic before the complementizer *ahte* was borrowed from some pre-Finnish variant or in East-Saamic before *šte/što* was borrowed from Russian. The systematic comparison with other Uralic languages might shed light on the linguistic reconstruction, but this endeavor was not in the scope of the present investigation.

It is possible that the relevant constructions were not overtly marked in Saamic languages prior to the contacts with languages of the Standard Average European type – i.e. Slavic and (the already Indo-Europeanized) Finnic. There is no evidence of any third etymological source for a *that*-type complementizer (beside *ahte* and *šte/što*), which might point towards this hypothesis.

The fact that *that*-type complementizers can, but need not be inserted in certain constructions such as direct quotations or embedded questions – yield-

ing asyndetic constructions – illustrates how overtly unmarked *that*-type complementation can function.

With respect to the evolution of complementation marking strategies towards the Standard Average European type, it is perhaps also worth noting that several of the contemporary Saami languages are in the process of borrowing complementizers once again – e.g. Skolt Saami *että* borrowed from Finnish is now a subvariant of *što* (as described in Section 2.2.) and Pite Saami *att* borrowed from Swedish is now the main variant replacing the original *ahte* (cf. Wilbur 2014: 245), even though this new development replaces formatives rather than changing the general syntactic structure.

The following sections briefly describe the diachronic sources of the complementizers under investigation. Although the etymology is not directly related to our analysis, it is worth mentioning that all complementizers discussed in this paper are loanwords. This can be seen from Table 1 in which complementizers of common descent are presented with the respective source languages.

Table 1: Canonical complementizers in Kildin, Skolt and North Saami presented as cognate sets

Kildin	Skolt	North	Etymology	Translation
		ahte	<Finnic	'that'
	että		<Finnish	'that'
šte	*što*		<Russian	'that'
	jos ~jõs	*jus ~jos ~juos*	<Finnic	'if'
koal'e			<Russian	'if'
jesl'e	*je´sli*		<Russian	'if'

4.1 *ahte*

According to Sammallahti (1998: 226), the North Saami subjunctor *ahte* 'that' is borrowed from Finnish *että*. Since the complementizer has functionally equivalent cognates in all other West-Saamic languages (cf. South, Pite and Lule Saami *ahte* 'that') it was probably not borrowed very recently into these languages individually, but goes back to an early form borrowed from Finnic at least into Common-West-Saamic.

4.2 *että*

The Skolt Saami subjunctor *että* 'that' is borrowed from contemporary Finnish *että* 'that' (see e.g. Feist 2010: 335). However, this marker is a marginal variant as compared to the much more widespread use of *što*, which is preferred in contemporary written Skolt Saami. We assume that *että* is a very recent borrowing characteristic only of the Skolt varieties spoken in Finland because it is neither attested in old recordings nor in the contemporary speech of Skolt Saami from the Russian side of the border.

4.3 *jesl'e/je´sli*

The subjunctors Kildin Saami *jesl'e* and Skolt Saami *je´sli* are borrowings from Russian *esli* (pronounced [jesl'i]) 'if', which goes back to a merged form of the copula *est'* [jest'] 'is' and the question particle *li* (see also Hansen, Letuchiy & Błaszczyk, this volume). The loan origin of these markers has been mentioned already by Itkonen (1958: 56). Note that the different spelling in the two languages reflects orthographic conventions rather than actual pronunciation. Note also that this marker is marginal in Skolt Saami, as it is used only by the speakers on the Russian side of the border.

4.4 *jos/juos*

The subjunctor North Saami *juos ~jus ~jos*, Skolt Saami *jõs ~jos* 'if' probably goes back to an early form borrowed from Finnic (cf. Finnish *jos* 'if') into West-Saamic languages and the two westernmost East-Saamic languages Inari and Skolt simultaneously (Korhonen 1981: 88, 94). Note that the different orthographic variants in both languages reflect pronunciations which are phonologically more or less adapted from the source word.

A more recent borrowing from Finnish is theoretically also possible (cf. Sammallahti 1998: 251 for North Saami), and the variant *jos* (pronounced [jos] as in contemporary Finnish) might have arisen due to recent Finnish influence.

4.5 *koal'e*

The subjunctor Kildin Saami *koal'e* is clearly borrowed from Russian. The source form is *ko-li* 'if, when' (an archaic equivalent of *esli*), although the time and place

of this borrowing are less clear. Itkonen (1958: 873), who does not propose an etymology, recorded the word only in the speech of one speaker from the Kīlltdialect in the 1940s. The widespread use of this form in today's written texts and its occurrence in the quasi-standard dictionary of Afanas'eva et al. (1985) – in which the variant *jesl'e* does not occur – makes *koal'e* the default form in contemporary Kildin Saami.

4.6 *šte/što*

The subjunctors Skolt Saami *što* and Kildin Saami *šte* are borrowed from Russian *čto* (earlier pronounced [čto], in contemporary language [što]) 'that' (see also Hansen, Letuchiy & Błaszczyk, this volume). This rather certain etymology was first mentioned by Itkonen (1958: 561). The different spelling of the vowel in the two languages reflects orthographic conventions rather the actual pronunciation. We have recorded several different variants in the speech of both Skolt and Kildin Saami speakers, such as [što] ~[štɒ] ~[štə]; in old sources even [čto] is attested (cf. Itkonen 1958: 561). Since the two other East-Saamic languages under Russian influence Akkala and Ter also use *što ~šte* (unlike Inari Saami *ät*, which is either cognate with the above mentioned North Saami *ahte* or borrowed from Finnish *että*) borrowing from Russian must have already taken place during some early stage of the development of the four easternmost Saami languages. The earliest known attestations are from the Bible translations, published in 1878 for Kildin and Akkala Saami (Matthew 1878) and in 1884 for Skolt Saami (Matthew 1884).

5 Summary

This chapter described the inventory of canonical complementizers and the syntax and semantics of related constructions in Kildin, North and Skolt Saami. We found relatively similar structures and functions regarding *that*- and *if*-type complementation in all three languages, albeit with conspicuous differences.

Kildin and Skolt Saami have one etymologically similar *that*-type complementizer: *šte/što*, which was borrowed from Russian. Whereas *that*-type complementation in Kildin Saami is formally contrasted to *if*-type complementation by using a different subjunctor (*koal'e*, which is also a Russian borrowing), in Skolt Saami, which uses *što* in both cases the semantic interpretation depends on the syntactic structure of the complement clause: *that*-complementation is formally licensed by *što* in combination with a declarative clause; *if*-complementation by

što in combination with a question. This results in the complementizer being used in addition to a formal question marker, i.e. either a question word (in the case of constituent questions) or a question enclitic (in the case of polar questions).

There is evidence that complementation in North Saami functions similarly to Skolt Saami, although traditional grammatical descriptions of North Saami characterize the language as exhibiting formal contrast between *that*- and *if*-type complementation by means of the two different subjunctors *ahte* and *jus*.

Our Kildin Saami data does not contain any combinations of the complementizer *šte* with constituent or polar question complements. Kildin Saami thus differs structurally from both Skolt and North Saami, in which the respective complementizers *što/ahte* are often combined with embedded questions. Based on these facts we described Kildin Saami as exhibiting a true *that*-type complementizer in contrast to Skolt and North Saami. In the latter languages, *that*-type is but one function of the respective multifunctional subjunctors.

References

Afanas'eva, Nina E., Aleksandra A. Antonova, Boris A. Gluchov, Lazer' D. Jakovlev & Ekaterina I. Mečkina. 1985. *Saamsko-russkij slovar'* [Saami–Russian dictionary]. Ed. by Rimma D. Kuruč. Moscow: Russkij jazyk.

Aikio, Ante. 2003. The geographical and sociolinguistic situation. In Jukka Pennanen & Klementti Näkkäläjärvi (eds.), *Siiddastallan. From Lapp communities to modern Sámi life* (Publication of the Inari Sámi Museum 5), 34–40. Jyväskylä: Siida Sámi Museum.

Bäuerle, Rainer & Thomas Ede Zimmermann. 1991. Fragesätze. In Armin von Stechow & Dieter Winderlich (eds.), *Semantik* (Handbücher zur Sprach- und Kommunikationswissenschaft 6), 333–348. Berlin & New York: de Gruyter.

Blokland, Rogier & Michael Rießler. 2011. Saami-Russian-Komi contacts on the Kola Peninsula. In Cornelius Hasselblatt, Peter Houtzagers & Remco van Pareren (eds.), *Language contact in times of globalization* (Studies in Slavic and General Linguistics 38), 5–26. Amsterdam & New York: Rodopi.

Ėndjukovskij, Aleksandr G. 1937. Saamskij (loparskij) jazyk [The Saami (Lapp) language]. In Georgij N. Prokof'ev (ed.), *Jazyki i pis'mennost' narodov Severa 1: Jazyki i pis'mennost' samoedskich i finno-ugorskich narodov* [Languages and literatures of the peoples of the North 1: Languages and literatures of the Samoyedic and Finno-Ugric peoples] (Naučno-Issledovatel'skaja Associacija Instituta Narodov Severa imeni P. G. Smidoviča. Trudy po lingvistike 1), 125–162. Moscow & Leningrad: Gosudarstvennoe učebno-pedagogičeskoe izdatel'stvo.

Feist, Tim. 2010. *A Grammar of Skolt Saami*. Manchester: University of Manchester PhD thesis.

Hansen, Björn, Alexander Letuchiy & Izabela Błaszczyk. This volume. Complementizers in Slavonic (Russian, Polish, and Bulgarian).

Itkonen, Toivo Immanuel. 1958. *Koltan- ja kuolanlapin sanakirja* [Skolt and Kola Saami dictionary] (Lexica Societatis Fenno-Ugricae 15). Helsinki: Suomalais-Ugrilainen Seura.

Kert, Georgij M. 1971. *Saamskij jazyk (kil'dinskij dialekt): Fonetika, morfologija, sintaksis* [The Saami language [Kildin dialect]: Phonetics, morphology, syntax]. Leningrad: Nauka.

Kert, Georgij M. 1986. *Slovar' saamsko-russkij i russko-saamskij* [Saami–Russian and Russian–Saami dictionary]. Leningrad: Prosveščenie.

Koponen, Eino, Jouni Moshnikoff & Satu Moshnikoff (eds.). 2010. *Sää´mǩiõll, ä´rbbǩiõl* [Saami language, heritage language]. Helsinki: Kotimaisten kielten tutkimuskeskus.

Korhonen, Mikko. 1981. *Johdatus lapin kielen historiaan* [Introduction to the history of the Saami language] (Suomalaisen Kirjallisuuden Seuran Toimituksia 370). Helsinki: Suomalaisen Kirjallisuuden Seura.

Matthew 1878 = *Mah'tveest Pas'-Evangeli. Samas* [The Gospel of Matthew. In Saami], translated by Arvid Genetz. Helsinki: British and Foreign Bible Society.

Matthew 1884 = *Gospoda mij Iisusa Christa Pas' Evangelie Matveest Same kille* [Our Lord Jesus Christ's Gospel of Matthew in Saami], translated by Konstantin Ščekol'din. Archangel'sk.

Moshnikoff, Satu, Jouni Moshnikoff & Eino Koponen. 2009. *Koltansaamen koulukielioppi* [A school grammar of Skolt Saami]. Inari: Sää´mte´ǧǧ.

Nickel, Klaus Peter. 1994. *Samisk grammatikk* [Saami grammar]. Oslo: Berlings.

Nielsen, Konrad. 1979. *Laerobok i lappisk (samisk). Utarbeidet på grunnlag av dialektene i Polmak, Karasjok og Kautokeino: Grammatikk. Lydlære, formlære, orddannelseslære og syntaks samt tillegg* [A textbook in Lapp (Saami). Based on the dialects in Polmak, Karasjok, and Kautokeino], 2nd edn. Oslo: Universitetsforlaget.

Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic descriptioni*, 2nd edn., 52–150. Cambridge: Cambridge University Press.

Nordström, Jackie & Kasper Boye. This volume. Complementizer semantics in the Germanic languages.

Rießler, Michael. 2009. Loanwords in Kildin Saami, a Uralic language of northern Europe. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages. A comparative textbook of lexical borrowing*, 384–413. Berlin: de Gruyter Mouton.

Sammallahti, Pekka. 1998. *The Saami languages*: *An introduction.* Kárášjohka: Davvi girji.

Sammallahti, Pekka. 2012. *Vuõ´lǧǧe jåa´tted ooudâs. Sää´mǩiõllsaž lookâmǩe´rjj* [Again we left. Skolt Saami text collection] (Publications of the Giellagas Institute 14). Oulu: University of Oulu.

Sammallahti, Pekka & Anastasija Chvorostuchina. 1991. *Unna sámi-saam', saam'- sámi sátnegirjjáš* [Short (North) Saami–(Kildin) Saami, (Kildin) Saami–(North) Saami dictionary]. Ohcejohka: Girjegiisá.

Sammallahti, Pekka & Klaus Peter Nickel. 2006. *Sámi-duiskka sátnegirj.* [Saami–German dictionary]. Kárášjohka: Davvi Girji.

Siegl, Florian & Michael Rießler. 2015. Uneven steps to literacy: The history of Dolgan, Forest Enets and Kola Sámi literary languages. In Heiko F. Marten, Michael Rießler, Janne Saarikivi & Reetta Toivanen (eds.), *Cultural and Linguistic Minorities in the Russian Federation and the European Union* (Multilingual Education 13), 189–230. Berlin: Springer.

Svonni, Mikael. 2009. *Samisk Grammatik* [Saami grammar], Kompendium. Tromsø: Universitetet i Tromsø.

Wilbur, Joshua. 2014. *A grammar of Pite Saami.* (Studies in Diversity Linguistics 5). Berlin: Language Science Press.

Gerson Klumpp

Semantic functions of complementizers in Permic languages

1 Introduction

1.1 The Permic languages

Permic (or Permian) is a branch of Uralic with two non-adjacent languages: Komi (Ko.) and Udmurt (Ud.) (see e.g. Riese 1998 and Rédei 1988 on Permic; Hausenberg 1998 on Komi, Csúcs 1998 on Udmurt). The core areas in which they are spoken lie within the European part of Russia, and they are covered by three standardized literary languages which relate to three administrative units of the Russian Federation: (i) *Komi-Zyrian* (KoZ) is the official titular minority language of the Komi Republic (capital Syktyvkar), protected and fostered by a language law dating back to 1992 (see e.g. Saarinen 2003). It had approximately 220,000 speakers in 2010. The northern variety of Komi-Zyrian, Ižma, is also spoken beyond the Komi Republic, in the Arkhangelsk Oblast and in Western Siberia. The Lower Vychegda dialect area was the base for the Old Zyrian (OZ) literacy tradition that began in the 14th century; (ii) *Komi-Permyak* (KoP) was the titular minority language of the Komi-Permyak Okrug (capital Kudymkar), which lost its official status as a minority district in 2005 and became a regular district of Perm Krai. It was spoken by approximately 63,000 speakers in 2010. Two more Komi varieties are spoken by small groups in the Vyatka territory and in the Jazva river area of Perm Krai. (iii) *Udmurt* (Ud.) is the acknowledged titular minority language of the Udmurt Republic (capital Izhevsk), but is also spoken in Tatarstan, Bashkortostan, as well as in the Vyatka Region. It counted approximately 340,000 speakers in 2010.

Lexical and structural influence from Russian is, of course, an issue for all three languages, and all have one or more registers characterized by the increased adoption of Russian patterns. It goes without saying that literary standards avoid these patterns. However, Russian influence on Komi is rather old, dating back to the time of the first contacts with Novgorod in the 11th century (see e.g. Leinonen 2002). Udmurt, on the other hand, has always been in more or less close contact with Turkic (Old Chuvash and Kipchak Turkic). In fact, it is the arrival of Turks in the area of the Permic proto-home in the Volga basin that was made responsible for the split of Proto-Permic into Komi and Udmurt (Belyh 2002). Russian-Udmurt contacts began only after the fall of Kazan under Russian rule in 1551, and they affected only the Northern Udmurt dialects. This is the background against

which one may state that, morphosyntactically, Udmurt represents more clearly the Altaic agglutinative type than Komi.

Other Altaic-type features characteristic for Udmurt are fairly rigid SOV, no clause initial subordinating conjunctions, and frequent use of converbs and nominalized predicates in clausal co- and subordination, while Komi has SVO as the prevailing word order, clause initial conjunctions, and less if any use of converbs and nominalized predicates in clause combining. While these differences correspond to contact areas and language development in historical times, it is less clear what the Permic proto-type looked like. In terms of word order, it is usually maintained that Udmurt-type SOV was also the basic word order in early Komi, whereas its later switch to SVO was due to Russian influence (Serebrennikov 1967: 101; Šutov 1999: 13). Concerning conjunctional syntax, there exists the view that the Finno-Ugric type consisted in juxtaposed finite predications, and either development, that towards Altaic type nominalization, as well as that towards Indo-European type conjunctional syntax, is secondary (Honti 1994) (see Section 2.2.1 on paratactic complements and the question of complementizer omission for a discussion).

Characterizing the phonological properties of the Permic languages, one has to mention their central vowels and abundance in alveolar and palatal sibilants (see below notes on transcription). Word stress is basically initial in Komi-Zyrian, free in Komi-Permyak and final in Udmurt (for a more detailed overview see Geisler 2005: 157–177). The morphological type is agglutinative, although without vowel harmony. Case systems are fairly rich, with 17 basic cases, plus additional local cases in Komi-Zyrian and Komi-Permyak. Permic languages have possessive declension, although in Komi, the functions of singular possessive suffixes became pragmatic. The development towards indicating 'givenness' resulted in the 2nd and 3rd person singular suffixes no longer indicating possessors (possessors are now exclusively expressed by possessive pronouns). For verbal morphology it should be noted that there is no object conjugation (object agreement on the verb), and the verbal mood system is rather poor: indicative and imperative, plus a conditional in Udmurt. The tense system, especially in Udmurt, is richer. Udmurt differentiates between present and future tense in all persons; Komi only in 3rd person. There are two past tenses, a simple past and a perfect; in addition there are several analytic past tenses in all languages which consist of a finite form in present, future or past tense and one of two predicative particles. The difference in meaning between the two particles parallels that between simple past and perfect, and includes notions of evidentiality, inferentiality, and mirativity (cf. Siegl 2004 and Leinonen 2000). Past tense formations operate along the parameters of anteriority, resultativity, habituality and evidentiality; in Udmurt there are at least seven (Winkler 2011: 97–101), in Komi eight (cf. Cypanov 2005)

past tenses. Sentential negation is carried out by a negative verb in all moods and tenses, except for the perfect in Southern Udmurt varieties, where we find a negative suffix on the perfect participle.

1.2 Data and transcription of examples

The data for this paper consists of examples taken from various (dialect) text publications (MF52, OU1–2, and W01 for Udmurt, and UK1–4, R78, and R96 for Komi), and from grammars and other studies (GSUJa. 1962, 1970, Karpova 1997, MSz., Šutov 1999, Jašina 1981 and Winkler 2011 for Udmurt, and L61, L52, KRS, Manova 1998, SKJa. II, and FF for Komi). Where possible, I present contrastive examples. These are taken from two parallel texts from the *Corpora of the research unit for Volgaic languages* (CVL) based at Turku University: *Pavlik Morozov* (PM) for Udmurt, Komi-Zyrian and Komi-Permyak, and *Suomi* for Udmurt and Komi-Zyrian. These parallel texts also have a Russian version. Finally, contrastive and other examples come from elicited data which were obtained in consultations (production and acceptability tasks) with two university educated native speakers of Udmurt and Komi at the University of Tartu.

In the case of Komi, it is not necessary to illustrate every phenomenon for Komi-Zyrian and for Komi-Permyak separately since the two languages are rather close. Differences are, of course mentioned and illustrated, but in many cases, a Komi-Zyrian or a Komi-Permyak example may serve to illustrate a certain pattern for entire Komi representatively.

All Komi and Udmurt examples are presented in a consistent transcription. Presenting examples from different sources in Cyrillic-based Komi and Udmurt orthographies, or in the (often narrow) phonetic transcription of the Finno-Ugric tradition would have burdened the reader considerably. The characters which may demand some explanation are the following: *ӧ* means an open central vowel, *ы* a closed one: both are illabial; in data from the Beserman dialect of Udmurt, both may fall together in [ə]; in addition, examples from Old Zyrian display the half closed vowels *ô* and *ê*, and in the Iźma dialect of Komi, as well as in dialects of Udmurt, secondary long vowels appear (e.g., *ē* < *ev*); palatality is indicated by the diacritic <’>, e.g., the palatal nasal by *ń*. Permic languages are rich in sibilants and affricates; cf. Table 1 for their graphic representation.

Table 1: Permic sibilants and affricates

	voiceless		voiced	
	sibilant	**affricate**	**sibilant**	**affricate**
alveolar	*s*	(*c*)[1]	*z*	–
palatal	*ś*	*ć*	*ź*	*ďž*
post-palatal	*š*	*č*	*ž*	*dž*

1.3 The content of the article

Section 1.4 introduces the main complement types and complementizers found in Permic. Section 2 offers a more detailed overview which lists all complementation strategies I could find in Permic languages, starting with finite complements. Non-finites follow, and finally attention is paid to some patterns which accompany complementation as well as to some marginal patterns. Section 2.2.1 on paratactic complements includes considerations on complementizer omission. Complementizer combinability is discussed in Section 3. It will be shown that different complementizers are not combined while retaining their original qualities. If two elements combine, only one can serve as a complementizer, while the other adds a semantic component that manipulates the fact type status of the complemented clause. Section 4 is about the opposition between finite factive and immediate perception complements on the one hand, and participial complements on the other. Participial subordination is a typical property of Permic (as well as of other Uralic and Altaic languages), and it is often considered a target for contact-influenced change towards finite subordination. However, both syntactic strategies co-exist, and it may be well asked what parameters trigger the preference of one over the other. In addition to arguments reported earlier, Section 4 argues for a discourse pragmatic opposition, i.e. participial complements contain preferably old information. Finally, Section 5 summarizes the results of the study.

The present contribution is not the first one treating complements (sentential actants) in Permic from a more general viewpoint (including semantic oppositions). In a recent paper by Serdobolskaja et al. (2012) these questions are also addressed. However, due to the fact that this paper has a wider perspective – including also Mari and Mordvin – it often does not illustrate phenomena with Permic examples.

1 This affricate is restricted to loanwords from Russian.

1.4 Complement types and complementizers

This section provides a general overview of the complementation in Permic languages, with references to the sections of the remainder of the paper where single phenomena are treated in more detail. It introduces semantic complement types ('semantic' according to the sentence mood) and their most important morpho-syntactic realizations. For an introductory example, see (1a)–(1c) from a parallel text in all three Permic literary languages. It contains a fact type (or declarative) complement of the complement taking predicate (CTP) 'know'. The complemented copula predication, 'I am Pasha's brother', is in brackets; in all three languages the bracketed string could function as an independent clause in exactly the same form. In addition, the complemented clauses contain a complementizer conjunction which takes the clause initial position in the Komi languages, but the clause final position in Udmurt. For the origin of these conjunctions, see Sections 2.1.1 and 2.1.2, respectively. Examples (2a) and (2b), containing also fact type complements, now with a CTP 'think', demonstrate for Komi and Udmurt that the use of a complementizer conjunction is in no way obligatory. Again, the complement clause could also function as an independent clause. See Section 2.2.1 on the question whether paratactic complements are historically primary.

(1) a. Komi-Zyrian (KoZ)
Sijö *e-z* *töd,* ***myj*** [*me* *Paša-lön* *vok*].
(s)he NEG.PST-3 know.CN **COMP** I Paša-GEN brother

b. Komi-Permyak (KoP)
Sija *e-z* *töd,* ***što*** [*me* *Paša-lön* *von*].
(s)he NEG.PST-3 know.CN **COMP** I Paša-GEN brother

c. Udmurt (Ud.)
So *u-g* *tody* *val,* [*mon* *Paša-len* *vyn-yz* *lu-iśko*]
(s)he NEG.PRS-3 know.CN was I Paša-GEN brother-3SG be-PRS.1SG
šuysa
COMP
'He didn't know that I am Pasha's brother.'
(PM: 1575)

(2) a. KoZ
Ćajt-a, [*skörm-as*].
think-PRS.1SG be.angry-FUT.3SG
'I think she'll be angry.'
(R96: 114.5)

b. Ud.
E, *ćakla-śko,* [*jyr-e* *byr-iz*].
INT think-PRS.1sg head-1SG get.lost-PST.3SG
'Hey, I think I lost my head.'
(OU1: 38.6)

Differently from (1) and (2), the complements in (3a) and (3b) could not function as independent predicates. They consist of a perfect participle, which functions as a direct object of the CTP 'see'. The participle is in the accusative case. The overt agent is expressed as a possessor in the ablative case and cross-referenced by a possessive suffix on the participle. The semantic type of the complement is not specified: it can be read as a fact type complement ('that') or as an immediate perception type complement ('how').

(3) a. KoZ
Oľga *aďď-is* [*Anna-lyś* ***kotört-öm-sö***] *da* *pyr* *žö* *mun-is*
Olga see-PST.3SG Anna-ABL **run-PP-ACC.3SG** and at.once also go-PST.3SG
bör-śań-ys.
after-EGR-3SG

b. Ud.
[*Anna-leś* *byźy-sa* ***pot-em-ze***] *aďďy-sa,* *Oľga* *so-len*
Anna-ABL run-CVB **go.out-PP-ACC.3SG** see-CVB Olga (s)he-GEN
börśaz *pot-iz*.
after.3SG go.out-PST.3SG

'Olga saw that/how Anna ran off and went after her immediately.' (elicited)

An explicit immediate perception type complement has to make use of the question word 'how' (KoZ *kyďži*, KoP *kyďž*, Ud. *kyźy*) as illustrated in (4a) and (4b). In Udmurt, the use of a participial complement and a 'how'-complement for the expression of immediate perception are mutually exclusive (see Section 3.3). 'How'-complements will be treated in more detail in Section 2.1.5, and participial complements in Section 2.3.2; Section 4 discusses the parameters complexity and information structure for the choice of a finite vs. non-finite complement.

(4) a. KoZ
Anna *aďďžyl-is,* [***kyďži*** *zonm-ys* *giž-is* *piśmö-sö*].
Anna watch-PST.3SG **how** boy-3SG write-PST.3SG letter-ACC.3SG

b. Ud.
Anna *śinjylt-iz,* [***kyźy*** *pijaš* *gožtet-se* *gožt-iz*].
Anna watch-PST.3SG **how** boy letter-ACC.3SG write-PST.3SG

'Anna watched how the boy wrote the letter.' (elicited)

Similar to participial complements, in Udmurt also predicative particles and locative expressions like *vań* 'there is', *kule* 'it's necessary', or *tatyn* '(be) there' may be complemented, as in (5); for more examples see Section 2.3.3.

(5) Ud. (elicited)
Anna [*gožtet-leś* *vań-ze*] *aďďž-iz*.
Anna letter-ABL there.is-ACC.3SG see-PST.3SG
'Anna saw there was a letter.'

Moving now to non-fact type (or non-declarative) complements, one must mention the following five: epistemic (or similative), conditional, imperative (or optative), interrogative, and conceptual (or activity) complements. The first group makes use of the borrowed Russian conjunction *budto(by)* 'as if', which occurs in Komi as *byťťö(kö)*, see (6), and in Udmurt registers with increased Russian influence as *butto(ke)* or *budto*, as in (7), or of the genuine Udmurt particle *kaď*, as in (8). In Russian, the status of this conjunction as a complementizer is not quite clear because it combines with the fact type complementizer *čto* (see Hansen, Letuchiy & Błaszczyk, this volume). The same problem holds for Permic languages where it combines, e.g., with the Komi *myj*. See Sections 2.1.7 and 3.2 for more examples and a short discussion.

(6) KoZ
Menym *kažiťć-ö* [***byťťö-kö*** *mam-ö* *menśim* *töd-ö*
I.DAT seem-PRS.3SG **COMP:EPIST** mom-1SG I.ABL know-PRS.3SG
kupajťć-öm-ös].
bath-PP-ACC.1SG
'It seems to me as if mother knew about my swimming.'
(Upper Vychegda; UK4: 196.28)

(7) Ud.
A*ś-me-len* *tańi* *tros-ez* *malpa-lo,* [***budto*** *eksej,*
self-1PL-GEN here many-3SG think-PRS.3PL **COMP:EPIST** tsar
kalyk-ez *žaľa-sa,* *syče* *manifest* *pott-iz*].
people-ACC pity-CVB such manifest take.out-PST.3SG
'Many of us here think that the tsar may have issued the manifest in feeling compassion with the people.'
(GSUJa. 1962: 331)

(8) Ud. (elicited)
[*Boris* *araky-ez* *u-g* *jaraty* *kad*] *pot-e.*
Boris vodka-ACC NEG.PRS-3 love.CN as.if come out-PRS.3SG
'It seems that (as if) Boris dislikes vodka.'

In addition, in Udmurt, there seems to be a strategy undescribed in literature, namely the modulation of the complementizer conjunction *šuysa* (lit. 'saying'). In (9), with the CTP 'dream', we find *kožasa* 'thinking' instead of *šuysa* in complementizer position. This marker may be interpreted as a non-fact equivalent to the fact type complementizer. However, this is the only instance of such modulation I came across, and the range of this strategy has not been investigated yet.

(9) Ud.
Vöta-j, *šu-em,* [*vu* *byr-em* *gurt-e* *vu-i* ***koža-sa***].
dream-PST.1SG say-PRF.3SG water end-PP village-ILL come-PST.1SG **think-CVB**
'I dreamt, she said, "thinking" I came into a village where the water had run dry.'
(W01: 89)

Conditional complements are not characterized by a special complementizer. In (10) the Udmurt variant in (10b) has the matrix verb in the conditional mood (in which the distinction of person is neutralized) and the complement clause is not specified for irrealis. The Komi variant in (10a), on the other hand, has an irrealis particle in the matrix clause, and the complement clause contains the conditional particle *kö*. This particle is enclitic and may not appear in clause initial position where it could take the function of a complementizer.

(10) a. Komi

Me	*eśkö*	*töd-i,*	[*völ-i*	***kö***	*Garri Potter-ys*	*menam*	*kerka-yn*].
I	IRR	know-PST.1SG	be-PST.3SG	**if**	Harry Potter-3SG	I.GEN	house-INE

b. Ud.

Mon	*tody-sal,*	[*Garri Potter*	*mynam*	*korka-n*	*val*	*šuysa*].
I	know-COND	Harry Potter	I.GEN	house-INE	has.been	COMP

'I would know if Harry Potter would have been at my house.' (elicited)

Imperative (optative) complements are characterized by the obligatory presence of the particle *med(ym)*, or its borrowed Russian equivalent *štoby* (cf. Hansen, Letuchiy & Błaszczyk, this volume), which carries the modal meaning (see Section 2.1.6 for more details). The complement verb is in future or past tense (the latter matches the Russian pattern). Udmurt and Komi differ in that in Udmurt the particle is integrated into the verbal paradigm where it forms together with a verb in future tense the optative mood, e.g., Ud. *med mynoz* '(s)he's supposed to go' (Winkler 2011: 103); note that according to the grammatical norm of Udmurt the imperative mood does not have 1st and 3rd person forms. In Komi grammar, there is no such normative decision, probably because here the particle does not necessarily appear in the immediate preverbal position. This, on the other hand, allows for the Komi particle to move to the initial position of the subordinate clause where it functions as a purposive or as a complementizer conjunction, as in (11a). In contrast, in Udmurt the particle has no conjunctional function, neither a function of an optative complementizer – cf. (11b), where the particle remains with the verb and the subordination is facultatively indicated by *šuysa* –, nor a purposive function (see Section 2.1.1 for the purposive function of Udmurt *šuysa*). The borrowed Russian particle *štoby* is illustrated in (12).

(11) a. KoZ

Anna	*kösj-ö,*	[***med***	*zonm-ys*	*giž-as*	*piśmö-sö*].
Anna	want-PRS.3SG	**COMP:OPT**	boy-3SG	write-FUT.3SG	letter-ACC.3SG

b. Ud.

Anna-len	*pot-e,*	[*pijaš*	*gožtet-se*	***med***	*gožt-oz*	*[šuysa]*].
Anna-GEN	go.out-PRS.3SG	boy	letter-ACC.3SG	**OPT**	write-FUT.3SG	COMP

'Anna wants that the boy writes the letter.' (elicited)

(12) KoP
Sećće *kaj-ys* *viśta-is* *ruć-ys-wö,* [***štoby*** *sija* *kywz-is*
then bird-3SG tell-PST.3SG fox-3SG-DAT **COMP:OPT** (s)he listen-PST.3SG
kaj-sö].
bird-ACC.3SG
'Then the bird told the fox to obey her.'
(Southern Permyak; R78: 474.6)

Interrogative complements contain WH-word questions or yes/no-questions. Considering the syntactic function of the question words within the first group, one may observe again that they function as complementizers in Komi, but not in Udmurt. This difference is illustrated by the contrastive examples in (13a)–(13c) where the question word 'why' functions as a linking element between embedding and embedded clause in both Komi languages ([13a] and [13b]), whereas the Udmurt version in (13c) has the question transformed into a participial complement in which the question word has no such linking function. While a structure as in (13c) is not possible in Komi, it is, however, possible in Udmurt to have a WH-word in pivot function as, e.g., illustrated in (14). See Section 2.1.4 on WH-word complements.

(13) a. KoZ
Deťinka *gögörvo-is,* [***myjla*** *vo-öma* *Śergej* *ďed-ys*].
boy understand-PST.3SG **why** come-PRF.3SG Sergei grandfather-3SG
b. KoP
Zonoćka *vežört-is,* [***myjlö*** *lokt-is* *Śerjoga* *ďed-ys*].
boy understand-PST.3SG **why** come-PST.3SG Sergei grandfather-3SG
c. Ud.
[*Śerjoga* *peśataj-leś* ***maly*** *lykt-em-ze*] *pijaš* *vala-z.*
Sergei grandfather-ABL **why** come-PP-ACC.3SG boy understand-PST.3SG
'The boy understood why Grandfather Sergej had come.'
(PM 903)

(14) Ud.
Ö-z *no* *vala,* [***kyź*** *gu-e* *uś-iz*].
not.PST-3SG and understand.CN **how** hole-ILL fall-PST.3SG
'And he didn't understand how he fell into the hole.'
(OU1: 65.4)

The second group of interrogative complements, yes/no-questions, behave quite similarly in Komi and Udmurt. There are two independent, but combinable strategies. One consists in the application of an enclitic question particle as in (15a) and (15b), where the particle occurs on the complement predicate; in Komi the verb is in clause initial position, in Udmurt in clause final, followed by the complementizer *šuysa*. The form of the interrogative complement clause does not differ from its independent equivalent, but is fully identical with it. The same holds for the

other strategy of yes/no-questions: the negative alternative. This pattern as well as combinations of both strategies are illustrated in Section 2.2.2.

(15) a. KoZ
*Anna jual-ö menśym, [giž-is-**ö** zonm-ys piśmö-sö].*
Anna ask-PRS.3S I.ABL.1SG write-PST.3SG-**Q** boy-3SG letter-ACC.3SG
b. Ud.
*Anna jua myneśtym, [pijaš gožtet-se gožt-iz-**a** šuysa].*
Anna ask.PRS.3SG I.ABL.1SG boy letter-ACC.3SG write-PST.3SG-**Q** COMP
'Anna asks me whether the boy wrote the letter.' (elicited)

Finally, conceptual (activity) complements, i.e. complements which do not express propositions, make use of the infinitive as, e.g., with 'think' in the Komi example (16). See Section 2.3.1 for infinitive complements in Komi and Udmurt.

(16) KoZ
Pop dumajt-is, [Jemöl'-ös vi-ni].
priest think-PST.3SG Jemöl-ACC kill-INF
'The priest thought (intended) to kill Yemel.'
(Upper Vyčeda; UK4: 6.30)

The complementation strategies addressed so far are of a more general character as they occur with different CTPs. However, from morphosyntactic or semantic points of view, there are more complement types to be mentioned. These types, which have a narrower distribution and are restricted to specific CTPs only, will be addressed in Section 2; among them, e.g., participle based essive complements and complements of the phasal CTP 'stop' or the CTP 'fear' in Section 2.3.2, Udmurt general converb complements with the phasal CTP 'finish' and the achievement CTP 'manage' in Section 2.3.4.

2 A survey of complementation strategies in Udmurt and Komi

This section provides a survey of those morphosyntactic strategies which occur with complement taking predicates in the Permic languages. Table 2 presents the structures to be discussed in this section and shows in which subsection they are treated. Besides the subordinating conjunctions, and the other finite strategies and non-finite verb forms treated in Sections 2.1–2.3, attention is also paid to some morphosyntactic phenomena which systematically co-occur with complementation, namely argument raising (Section 2.4), quotative particles and reference strategy (Section 2.5), as well as some peripheral strategies which happen to

occur with CTPs (Section 2.6). The survey includes other functions of complementizer conjunctions and complementation markers, as well as etymological information about the elements in question. In general, special attention is paid to differences between Udmurt and Komi, and, when necessary, between Komi-Zyrian and Komi-Permyak. The question of combinability of complementation markers is separately treated in Section 3. Exact information about the distribution of each complementation marker with different classes of CTPs, unfortunately, cannot be provided in this survey, except for some minor strategies which are restricted to specific complement taking predicates. All these questions, however, cannot be equally addressed throughout this survey, and, in parts, a somehow impressionistic discussion is put up with.

Table 2: Permic complementation strategies treated in Section 2.

subordinating conjunctions	other finite clause strategies	non-finite verb forms
Ud. *šuysa* 'saying' (2.1.1) Ko. *myj*, Ko., Ud. *što* 'what' (2.1.2) OZ *veveś* (2.1.3) WH-words (2.1.4) Ko. *kyďž(i)*, Ud. *kyźy* 'how' (2.1.5) Ko. *med(ym)*, *medby*, Ko., Ud. *štoby* 'that, in order to' (2.1.6) Ko. *byťťökö*, Ud. *budto* 'as, as if' (2.1.7)	parataxis (2.2.1) question particle (2.2.2) negative alternative (2.2.2) conditional particle (2.2.3)	infinitive (2.3.1) participles (2.3.2) predicative particles (2.3.3) converbs (2.3.4)

2.1 Complementizers based on subordinating conjunctions and question words

2.1.1 *Udmurt* šuysa

Udmurt *šuysa* (dial. *šusa*) is a general converb in *-sa* of the verb *šuy-* 'say, call'. It has acquired a couple of functions, among them that of a clause final indicator of subordination which does also occur with complement clauses as was already illustrated in (1c), (10b), (11b), and (15b) above. Šutov (1999: 28) calls it along with *bere* 'once', *dyrja* 'while', and *vylyś* 'though' a "postpositional conjunction" (*postpozitivnyj sojuz*), because the conjunction always terminates a subordinated clause, meaning (i) 'that (Russ. *čto*)', (ii) 'that, in order to (Russ. *čtoby*)'. Analogical formations are found in Mari (*manən*), Turkic (Tatar *dip*, Chuvash *tese*) and Mongolic languages; in case of Udmurt, the responsible influence comes

from Tatar, i.e. *šuy-sa* 'say-CVB' is a calque from Tatar *di-p* 'say-CVB' (cf. Bartens 2000: 311); note, however, Šutov's (1999: 29–30) proposal for parallel developments due to identical word order conditions in these languages. A concise overview of the functions of *šuysa* is presented by Jašina (1981), and earlier by Perevoščikov (1959: 246–248); Šutov (1999: 38–39) investigates the use of *šuysa* in Udmurt Gospels from mid 19th century. Recently a parallel grammaticalization path for Udmurt *šuysa* and Mari *manən* has been sketched by Serdobolskaja and Toldova (2011).

There are four types of subordinate clauses *šuysa* appears in: complement clauses, as in (17a), purposive clauses, as in (17b), causal clauses, as in (17c), and circumstantial or chaining clauses, as in (17d). In addition, there are two "pseudo-adnominal", or postposition-like uses (see Kel'makov 1973: 127–130); one comes with verbs of naming, meaning roughly 'as'; see (18a). This function is in competition with a construction with the instrumental case. The other postposition-like use comes with dative marked goals as an equivalent to the postposition *ponna* 'for', yielding, e.g., a consecutive object as in (18b).

(17) Conjunction functions of Udmurt *šuysa* (elicited)

a. Clause final conjunction in a complement clause:
Anna [*pijaš* *gožtet* *med* *gožt-oz* ***šuy-sa***] *malpa-z.*
Anna boy letter OPT write-FUT.3SG **say-CVB** think-PST.3SG
'Anna thought that the boy should write a letter.'

b. Clause final conjunction in a purposive clause:
Pijaš *gožtet* *med* *gožt-oz* ***šuy-sa***, *Anna* *so-ly* *bumaga*
boy letter OPT write-FUT.3SG **say-CVB** Anna (s)he-DAT paper
śot-iz.
give-PST.3SG
'In order for the boy to write a letter Anna gave him paper.'

c. Clause final conjunction in causal clause:
Pijaš *gožtet* *gožt-iz* ***šuy-sa***, *Anna* *so-ly* *kuźym*
boy letter write-PST.SG **say-CVB** Anna (s)he-DAT gift
baśtiz.
buy-PST.3SG
'Because the boy has written a letter Anna bought him a gift (as a reward).'

d. Chaining converb (original function):
Anna *"pijaš* *gožtet* *med* *gožt-oz"* ***šuy-sa*** *pot-iz.*
Anna boy letter OPT write-FUT.3SG **say-CVB** go.out-PST.3SG
Saying "the boy should write a letter" Anna went out.

(18) a. Pseudo adnominal function of the Udmurt š*uysa* with verbs of designation:
So *dyry-śen* *so* *ńur* *inty-jez* *Bodjo* *šur* ***šu-sa*** *ńima-ľľam.*
this time-EGR this swamp place-ACC Bodjo river **say-CVB** call-PRF.PL
'Since that time they call this swampy place Bodjo river.'
(OU2: 248.9)

b. Pseudo adnominal function of Udmurt *šuysa* with dative marked goals:
gubi-ly ***šuy-sa*** (~ *gubi* *ponna*) *pot-em* *val*
mushroom-DAT **say-CVB** mushroom for go.out-PP was
'(s)he had gone for mushrooms'
(Kel'makov 1973: 130)

As a complementizer, *šuysa* occurs with all kinds of CTPs: utterance, propositional attitude, commentative, and knowledge predicates, but not in immediate perception contexts. It is often not obligatory. Winkler (2011: 169) observes that with *verba dicendi et sentiendi* the complementizer conjunction is often absent, and the criteria for its use are often not grammatical but "contextual", e.g., if the complement clause is very short, as in (19), *šuysa* must be used.

(19) Ud.
So *vera-z* [*myn-iśko* ***šuysa***].
(s)he say-PST.3SG go-PRS1SG **COMP**
'He said that he is coming.'
(Winkler 2011: 169)

Winkler observes also that the complementizer conjunction is superfluous if the matrix CTP is a finite form of the verb *šuy-* 'say', and the use of the conjunction would result in a *figura etymologica*. Indeed, as a survey of narrative folklore texts (Wichmann 1901) has shown, reported speech is complemented by *šuysa* if the utterance CTP is *vera-* 'say, tell', *jua-* 'ask', *ćyrekty-* 'shout', *kargy-* 'curse' etc., but not if it is *šuy-* 'say'. This observation was confirmed by a native speaker who refrains from using the complementizer *šuysa* with this CTP, cf. (20).

(20) Ud. (elicited)
Anna *vera /* *ʔšu-e,* [*pijaš* *gožtet-se* *gožt-e* ***šuysa***].
Anna say.PRS.3SG say-PRS.3SG boy letter-ACC.3SG write-PRS.3SG **COMP**
'Anna says that the boy writes the letter.'

Vilkuna (1998: 185) sees interdependency between the use of *šuysa* and the word order. According to her, the complementizer is obligatory when the object complement takes the O position in a SOV structure as in (21a), whereas it is facultative when the complement clause follows the matrix clause, as in (21b).

(21) Ud.
a. *Nina* [*mon* *ćoryg* *śi-iśko* ***šuysa***] *malpa-z.*
Nina I fish eat-PRS.1SG **COMP** think-PST.3SG
b. *Nina* *malpa-z,* [*mon* *ćoryg* *śi-iśko* *[**šuysa**]*].
Nina think-PST.3SG I fish eat-PRS.1SG **COMP**
'Nina thought that I eat fish.'
(Vilkuna 1998: 185)

Modern Udmurt orthography puts a comma after the sentence final conjunction if the subordinate clause precedes the main clause. Sometimes, however, authors feel that the *šuysa*-clause is still an adverb and should not be separated from the main verb by this orthographic device (cf. Šutov 1999: 18).

The complementizer conjunction *šuysa* combines freely with interrogative or modal markers and is thus not restricted to a specific semantic type of complements. Winkler (2011: 166) calls it a facultative indicator of subordination. Combinations of *šuysa* with WH words will be addressed in Section 3.2 and the combination of *šuysa* and the borrowed *što* in Section 3.1.

2.1.2 The fact type complementizer 'what'

'What'-complements correspond to a subgroup of English *that*-complements. In Komi literary language, as well as in Udmurt Russian-influenced substandard varieties, there is a complementizer conjunction meaning lit. 'what', which precedes the complemented sentence. It introduces subject as well as object sentences. In Komi-Zyrian, this conjunction is the genuine interrogative word *myj* (dial. *möj*) 'what', a calque of the Russian *čto* '1. what; 2. that (complementizer)'. In dialects and substandard varieties the Russian conjunction appears as a loanword, e.g., Luza-Letka *ištö* (Žilina 1985: 150), Udora *štö*, *yštö*, *uštö* (Sorvačëva & Beznosikova 1990: 90) (see also Popova 2011b). The same holds for Udmurt dialects and substandard. In Komi-Permyak the loan *što* (dial. *štö*) has been codified in the literary language. In substandard the two conjunctions appear also in pleonastic combinations where an older genuine and a new borrowed element fulfill the same function; Žilina (1985: 112), for example, reports the occurrence of *štö myj* in the Luza-Letka dialect.

It is commonly assumed that complementizer conjunctions are prevalent in Komi but not in Udmurt, due to the earlier influence of Russian on Komi (Šutov 1999: 18; Manova 1998: 170); cf. the Komi sentences in (1a) and (1b), repeated here as (22a) and (22b). (23) is a Komi-Zyrian dialect example of the borrowed Russian conjunction, and (24) of the combination of the genuine and the borrowed conjunction.

(22) a. KoZ

Sijö	*e-z*	*töd,*	[***myj***	*me*	*Paša-lön*	*vok*]
(s)he	NEG.PST-3	know	**COMP**	I	Paša-GEN	brother

b. KoP
Sija *e-z* *töd,* [***što*** *me* *Paša-lön* *von*]
(s)he NEG.PST-3 know **COMP** I Paša-GEN brother
'He didn't know that I am Pasha's brother.'
(PM 1575)

(23) KoZ
Seki *byd* *ńegramötnej-ys* *kaźal-as,* [***štö*** *o-s* *poź* *mun-ny*
then every illiterate-3SG notice-FUT3SG **COMP** NEG.PRS-3SG can go-INF
mat-ö].
near-ILL
'Then every illiterate will notice that one must not approach.'
(Udora; UK1: 354.13)

(24) KoZ
Me *i-g* *dumajt,* [***yštö-myj*** *zerm-as*] (...)
I NEG-PST.1 think-CN **COMP** rain-FUT3SG
'I didn't think that it will rain (...).'
(Vym; Žilina 1998: 64; Popova 2011b: 72)

Concerning word order, the complementizer conjunction in Komi usually takes the initial position in the complemented sentence, but this is not obligatory, as the inversion of subject and conjunction in the following example from Komi literary language shows.

(25) KoZ
Tačömtor-sö *vo* *kyź* *saj-yn* *ńekod* *e-z* *ćajtly* *lo-ig,*
such.thing-ACC.3SG year 20 ago-INE nobody not.PST-3 suspect.CN be-CVB
mövpal-öny *völi,* [*perym* *gižöd* *kyv-jas* ***myj*** *o-z*
think-PRS.3PL was Permic written language-PL **COMP** not.FUT-3
vežśy-ny...],
change-CN.PL
'Nobody would have expected such a thing twenty years ago, [then] they thought that the Permic literary languages would not change (...)'
(Cypanov 2010: 32)

Komi *myj* appears also as conjunction with at least two other meanings which should be mentioned and illustrated here: in (26a) we find it as a sentence connector with the meaning 'so that' (with an inferential reading of the perfect tense), and in (26b) as a causal conjunction, probably shortened from *sy ponda*, *myj* 'because (lit. 'for this, that...').

(26) KoZ

a. Upper Vyčegda

I	*ńoľ*	*peľes-in*	*pyzan*	*vyl-in*	*ly*	*ćukar,*	***myj***	*tani*
and	four	corner-INE	table	on-INE	bone	pile	**what**	here
śoj-öma-öś	*noľ-en.*							
eat-PRF.3-PL	four-INS							

'And at the four corners of the table there are bone piles so that there must have eaten four persons.'
(UK4: 74.31)

b. Upper Vyčegda

A	*Śemuk-Śemö-ös*	*bör*	*leďž-isny*	*gort-as,*	***myj***
but	Semuk-Seme-ACC	back	let-PST.3PL	home-ILL3SG	**what**
Śemuk-Śemö-lyś	*e-z*	*aďďži-ny*	*ńe-kučöm*	*myž.*	
Semuk-Seme-ABL	not.PST-3	see.CN-PL	no-what.kind	guilt	

'But Semuk-Seme was released back home, because they couldn't find any guilt with Semuk-Seme.'
(UK4: 76.26)

The borrowed complementizer conjunction *što* may also appear in Udmurt substandard where it either combines with sentence final *šuysa* (see Section 3.1), or comes alone as, e.g., in (27). The opposition between the genuine Udmurt complementizer and the borrowed *što* is usually not considered to be a functional one in the sense that we have different semantic types of complements, but rather one of language change under language contact (see the "framing construction" in Section 3.1), and it is assumed that they appear in synonymous constructions. However, the following observation from the Beserman dialect of Udmurt is interesting as it seems to display a functional reason for the choice of the Russian complementizer conjunction. As Biryuk and Usačeva (2011: 263) report "there is a large amount of Beserman constructions, which in our corpus (and according to our experience) are never replaced with Russian ones. Nevertheless, cases of substitution do occur [...]". It can be observed that the sentence in (27) contains not only one but two complement predications: (i) 'girl feeds', (ii) 'one chicken is missing'. One may assume that the Russian model with the clause initial complementizer *što* is more convenient for this sequence of complements than the sentence final *šuysa*. However, the details are a matter of future research.

(27) Ud.

aďď-iz	[***što***	[*nəlmurt*	*śud-e*]		[*odig-ez*	*petuh-ez*	*evəl*]],
see-PST.3SG	**COMP**	girl	feed-PRS.3SG		one-3SG	chicken	is.not
pići	*pi*	*petuh*	*punna*	*mən-iz*	*biźə-sa.*		
little	boy	chicken	for	go-PST.3SG	run-CVB		

'[When] he saw that the girl was feeding (birds) and that one chicken was missing, the boy ran for the chicken.'
(Beserman, Biryuk & Usačeva 2011: 263)

2.1.3 The Old Zyrian complementizer *veveś*

The Old Zyrian literary language uses the conjunction *veveś* as an equivalent of the modern Komi what-complementizers *myj* and *što*. In the Old Zyrian corpus this conjunction appears eight times altogether; in six cases it functions as a complementizer, with three different CTPs, once with 'believe', see (28a), four times with 'say' as in (28b), and once with 'be amazed'. In two instances, it can be read as causal conjunction, as in (29). The exact functional range of this word is unclear. Etymologically, it is a reduplication of *veś* (L52: 124), a functional word which has a cognate in Mansi (*vos* 'in order to, that') and appears also in other compound forms as, e.g., in the irrealis particles *veśkö*, *veśig* (L52: 124; KÉSK 1999: 53–54).

(28) Old Zyrian complementizer *veveś*

a. *Esk-a-mnym kö* [***veveś*** *iisus kul-y*] (...)
believe-FUT-1PL if **COMP** Jesus die-PST
'For if we believe that Jesus died (...)'.
(L52: 65, lines 15–16 [1 Thessalonians 4, 14])

b. *Ina veśtas vöjp-a-m tijan-ly,* [***veveś*** *lokt-ö jezi*] (...)
true right say-FUT-1SG ye-DAT **COMP** come-PRS time
'Verily, I say unto you, the hour is coming (...).'
(L52: 66, lines 51–53 [John 5, 25])

(29) Old Zyrian

jordǯal-am-ö menam vim veś; ***veveś*** *o-g korśy as gaž-ös*
judge-PP-1SG I.GEN COP just **because** NEG.PRS-1SG search.CN own will-ACC
'(...) and my judgement is just; because I seek not mine own will.'
(L52: 67, lines 76–79 [John 5, 30])

2.1.4 Question word complements

Cross-linguistically, complements are often introduced by question words which function as a pivot element between the matrix sentence and the embedded sentence. This is true also for the Permic languages. Two WH-words 'what' and 'how' also function as complementizers on a broader level (see Section 2.1.2 above, and Section 2.1.5 below). In Komi grammar, this double function is well-known (cf. SKJa. II: 212 for comparison of WH and fact type complements); consider the minimal pair in (30).

(30) KoZ (elicited)

a. *Menym viśtal-isny,* [***myj*** *sijö vaj-is*].
I.DAT tell-PST.3PL **what** (s)he bring-PST.3SG
'They told me what he brought.'

b. *Menym viśtal-isny,* [***myj*** *sijö vaj-is ńań*].
I.DAT tell-PST.3PL, **COMP** (s)he bring-PST.3SG bread
'They told me that he brought bread.'

In Udmurt grammar (GSUJa. 1962: 333), WH-words are included in the class of "conjunctional words", because, just like conjunctions, they express a conjunction of a subordinate and a main clause, but are different from the latter, as they are at the same time also sentence constituents (GSUJa. 1962: 333). In general, WH-complements occur with all kinds of CTPs, e.g., with the knowledge predicate 'know' in (31). The Udmurt complementizer *šuysa* may combine with WH-complements (see Section 3.2); in Komi this combination is less accepted.[2]

(31) a. Komi
Me kö töd-i, [***kytyn*** *sijö yčk-ö*], *otsal-i eśkö.*
I if know-PST1SG **where** (s)he mow-PRS.3SG help-PST.1SG MOD
b. Ud.
Mon tody-sal ke, [***kytyn*** *so turna*], *jurtty-sal.*
I know-COND if **where** (s)he mow-PRS.3SG help-COND
'If I had known where she was mowing, I would have helped (her).' (elicited)

The status as pivot element between embedding and complement clause is, in principal, different in Komi and Udmurt since in Udmurt the WH-word remains in its original position (cf. Šutov 1999: 20); in general, Udmurt WH-words do not have to take the clause-initial position (cf. *könja* in *Aďami könja syle?* 'How much does a person cost', Šutov 1999: 19), and they keep their position also in non-finite constructions. In Komi, on the other hand, if the WH word does not take the clause-initial position, this is due to functional changes of word order as, e.g., in the case of topicalization; cf. (32).

(32) KoZ
Sy din-ö vo-an da juav sy-lyś, [*med ďžoľa ćoj*
(s)he at-ILL come-PRS.2SG and ask.IMP.2SG (s)he-ABL SPL young sister
kön *ol-ö*].
where live-PRS.3SG
'(When) you come to her you ask her where she lives, the youngest sister.'
(Upper Vychegda; UK4: 12.29)

2 The combination of the two different *myj*-functions from (30) is considered ungrammatical by the Komi consultant; cf.:

(i) KoZ (elicited)
**Menym viśtal-isny,* [*myj myj sijö vaj-is*].
I.DAT tell-PST.3PL CC what (s)he bring-PST.3SG
Intended: 'They told me what he brought.'

2.1.5 The immediate perception complementizer ‘how’

Besides ‘what’ (see Section 2.1.2 above) another complementizer based on a question word is Komi-Zyrian *kyďž(i)* ‘how; as’, Komi-Permyak *kyďž* ‘how; as’, and Udmurt *kyźy* ‘how; as’, equivalents of English *how* or Russian *kak*. A clear instance of an immediate perception complement with CTP ‘see’ is illustrated in (33). Permic languages may display aspect sensitivity with these complement clauses. This is illustrated for Komi in (34): the fact type complement in (34a) has a simple CTP ‘see’ and the complement predicate is in perfect tense. Differently, the ‘how’-complement in (34b) needs a verb of imperfective aspect which is derived from the perfective one with the help of the frequentative suffix *-l-*; in addition, the tense of the complement predicated is not perfect, but “first past” (or preterite). In other words, *myi* and *kyďži* are semantically contrasted in the sense that they are licensed by different aspects of the matrix verb and they license different tenses of the complement verb.

(33) a. Komi
*Anna vidžöd-ö, [**kyďži** zonm-ys giž-ö piśmö-sö].*
Anna watch-PRS.3SG **how** boy-3SG write-PRS.3SG letter-ACC.3SG

b. Ud.
*Anna śinjylt-e, [**kyźy** pijaš gožtet-se gožt-e].*
Anna watch-PRS.3SG **how** boy letter-ACC.3SG write-PRS.3SG

‘Anna watches how the boy writes the letter.’ (elicited)

(34) KoZ (elicited)

a. *Anna adďž-is, [**myj** zonm-ys giž-öma piśmö-sö].*
Anna see-PST.3SG **COMP** boy-3SG write-PRF.3SG letter-ACC.SG
‘Anna saw that the boy wrote the letter.’

b. *Anna adďžy-l-is, [**kyďži** zonm-ys giž-is piśmö-sö].*
Anna see-FRQ-PST.3SG **how** boy-3SG write-PST.3SG letter-ACC.3SG
‘Anna saw how the boy wrote the letter.’

In Komi, *kyďži* ‘how; as’ functions also as a comparative correlative conjunction (‘as’), see (35a), as an “essive preposition”, see (35b), and in several dialects as a temporal (‘when’), see (35c), or a causal conjunction, see (35d).

(35) a. KoZ
*Aďďž-isny, myj sijö siďž i em, **kyďži** i viśtal-isny.*
see-PST.3PL COMP it so also is **how** also say-PST.3PL
‘They saw that it was so as it had been told.’
(Upper Sysola; UK3: 42.23)

b. KoZ
Baťe *kut-is,* ***kyďži*** *ydžyd* *mort,* *mijanes* *veled-ni*
father begin-PST.3SG **as** big man we.ACC teach-INF
'Father, as the boss, began to teach us.'
(Upper Vychegda; UK IV: 410.3)

c. KoP
Toľko *möd-is* *kaj-ny* *gortas,* ***kyďž*** *kaźa-is* *ośta-as*
only begin-PST.3SG go-INF home **as** notice-PST.3SG ice.hole-INE.3SG
šuka-ös.
pike-ACC
'He was just about to start home when he noticed a pike in the ice-hole.'
(Southern Permyak; R78: 492.3)

d. KoZ
A *Ivan* *bara* *o-z* *götraś,* ***kyďži*** *abu* *polnej* *um-a.*
and Ivan again not.PRS-3 marry.CN **as** is.not full mind-ADJ
'But Ivan does not get married since he is a bit cracked.'
(Upper Vychegda; UK IV: 274.14)

2.1.6 The optative complementizers *med(ym)*, *medby*, and *štoby*

The difference between a fact type and an optative complementizer, e.g. Latin *quod* vs. *ut*, or Russian *čto* vs. *čtoby*, is also typical for Komi as well as for Udmurt substandard (i.e. for registers with increased Russian influence). In Standard Udmurt, on the other hand, *šuysa* is used in both contexts (cf. Šutov 1999: 17). But in both Permic languages the complementizer conjunction which is used in optative complements is the same as in purposive clauses; cf. the Udmurt example (17b) above. In Komi, the genuine optative complementizer conjunction is the particle *med* 'in order to, that, shall be' (< Proto Permic **mẹd*; KESK: 17), or *medym*; in contexts of negation the particle may occur in contracted form, e.g., *moz* < *med oz* (*o-z* = present tense 3rd person singular of the negative auxiliary). In Udmurt the contracted negative forms are less transparent, e.g., in the 3rd person singular *medaz* the segment *-az* does not correspond to the 3rd person singular present tense of the negative auxiliary, *ö-z*, or the future tense, *u-z*.[3] The Udmurt particle is sometimes considered to be part of the verbal paradigm, namely as a formative of the optative mood (e.g. Winkler 2011: 103). For Komi such a view does not exist. The main difference between Komi and Udmurt is the position of the particle in a subordinated clause: in Udmurt it always remains in preverbal position, but in Komi it moves to a clause-initial position; cf. (36). The Komi particle is thus an example of grammaticalization of a modal particle into conjunction

3 Serdobol'skaja et al. (2012: 430) report for the Beserman dialect of Udmurt the generalized use of the 1st person singular *medam* in all persons.

(Manova 1975). Šutov (1999: 46–47) claims that, just like in Komi, the particles *med, medaz* etc. can function also in Udmurt as purposive conjunctions, but in all his examples the particle remains in the preverbal slot, and does not move to the beginning of the sentence, like in Komi or Russian, or to the end of the sentence, like the genuine conjunction *šuysa* (see also the end of this section).

(36) a. Komi
*Anna kösj-ö / Kol-ö [**med** zonm-ys*
Anna want-PRS.3SG be.necessary-PRS.3SG **COMP:OPT** boy-3SG
giž-as piśmö-sö].
write-FUT.3SG letter-ACC.3SG

b. Ud.
*Anna-len pot-e / Kul-e [pijaš gožtet-se **med***
Anna-GEN go.out-PRS3SG be.necessary-PRS3SG boy letter-ACC3SG **OPT**
gošt-oz [šuysa]].
write-FUT.3SG COMP

'Anna wants/It's necessary that the boy writes the letter.' (elicited)

In an overview of the functions of the particle *med* in Permic, Popova (2011a) lists beside the optative particle/complementizer function and the purposive conjunction, the prefixal use in superlative formations: e.g. Ko. *med-jon* 'the strongest' (*jon* 'strong'). Besides *med(ym)*, the Russian loan *štoby* (and variants of it) fulfils the same function, especially in Komi Permyak, cf. (37). In addition there is also the hybrid form *medby*, found in Zyrian as well as in Permyak varieties of Komi; see (38). As already said, the Russian loan occurs also in Udmurt registers of increased Russian influence. Furthermore, Udmurt has an alternative structure with the necessitive predicative participle in *-ono*; see Section 2.3.4 below. The functional equivalent of *med(ym)/štoby* in the Komi-Jazva dialect (KYaD: 83), and partly also in Permyak (Cypanov 2005: 54), is *aś*.

(37) KoP
*Sećće kaj-ys viśta-is ruć-ys-wö, [**štoby** sija kywz-is*
then bird-3SG tell-PST.3SG fox-3SG-DAT **COMP:OPT** (s)he listen-PST.3SG
kaj-sö].
bird-ACC.3SG
'Then the bird told the fox to obey her.'
(Southern Permyak; R78: 474.6)

(38) KoP
*Vot i pond-is kelmiś-ny starik Ivan-lö, [**medby** sija*
well and begin-PST.3SG beg-INF old.man Ivan-DAT **COMP:OPT** (s)he
dugd-is ors-ny.
stop-PST.3SG play-INF
'The old man began to beg Ivan to stop playing the accordion.'
(Northern Permyak; R78: 424.19)

With utterance verbs like 'tell', which can either have a stating or a manipulating meaning the optative complementizer stands in opposition with the 'what' complementizer; cf. (39a) and (39b) from Komi-Zyrian with *juört-* 'tell'; (40) illustrates the same contrast in Komi-Permyak. The tense selection after the optative complementizer is restricted, as is demonstrated in (41a) and (41b). However, under the influence of Russian the optative complement predicate may also come in past tense in some varieties of Komi. Serdobol'skaja et al. (2012: 432–433) report such a development for the Ižma as well as for the Pečora dialect.

(39) KoZ

a. *Juört-isny* *car-ly,* [***möj*** *ićöt* *soć-lön* *röďťć-is* *ćört*
tell-PST.3PL tsar-DAT **COMP** little sister-GEN be.born-PST.3SG devil
koď *pi*].
like boy
'They informed the tsar that the little sister had given birth to a devil-like boy.'

b. *Car* *skörm-is* *da* *i* *juört-is* *soć-jas-ly,* [***med***
tsar get.angry-PST.3SG and also tell-PST.3SG sister-PL-DAT **COMP:OPT**
pi-sö *lovjön* *šuj-isny* *böćka-ö* *da* *leďž-isny* *more-ö*].
boy-ACC.3SG alive put.in-PST.3PL barrel-ILL and let-PST.3PL sea-ILL
'The tsar got angry and told the sisters that they should put the boy in a barrel alive and set him out into the sea.'

(Luza-Letka; UK3: 6.20–22)

(40) KoP

a. *Kaj-ys* *višta-is* *ruć-ys-wö,* [***što*** *sija* *verd-as* *ruć-sö*].
bird-3SG say-PST.3SG fox-3SG-DAT **COMP** (s)he feed-FUT.3SG fox-ACC.3SG
'The bird said to the fox that it will feed the fox.'

b. *Sećće* *kajys* *višta-is* *ruć-ys-wö,* [***štoby*** *sija* *kywz-is*
then bird.3SG say-PST.3SG fox-3SG-DAT **COMP:OPT** (s)he listen-PST.3SG
kaj-sö].
bird-ACC.3SG
'Then the bird told the fox to obey her.'

(Southern Permyak; R78: 474.4–6)

(41) KoZ (elicited)

a. *Marko* *šu-is,* [***myj*** *sijö* *dugd-as* / *dugd-ö* / *dugd-is* /
Marko say-PST.3SG **COMP** (s)he stop-FUT.3SG stop-PRS.3G stop-PST.3SG
dugd-öma *vors-ny*].
stop-PRF.3SG play-INF
'Marko said that he will stop/stops/stopped/has stopped playing.'

b. *Marko* *šu-is,* [***med[ym]*** *sijö* *dugd-as* / **dugd-ö* /
Marko say-PST.3SG **COMP:OPT** (s)he stop-FUT.3SG stop-PRS.3G
**dugd-is* / **dugd-öma* *vors-ny*].
stop-PST.3SG stop-PRF.3SG play-INF
'Marko said he should stop playing.'

The optative complementizer in Komi is also used to indicate irrealis contexts, as with the CTP 'remember' in (42).

(42) KoZ
O-z *pomńit* [***med*** *tačöm* *ćöskyda-sö*
NEG.PRS-3 remember.CN **COMP:OPT** such delicious-ACC.3SG
kor-kö *śojl-is*]
when-INDF eat-PST.3SG
'He does not remember that he should have ever eaten something delicious like this.'
(Popova 2011a: 200–201)

The double function of Komi *med* and its variants as optative complementizers and as conjunctions for purposive adjuncts is illustrated in (43) from Old Zyrian, and in (44), from 20th century Komi varieties. In (43), the first occurence of *med* is a complementizer (*med tödadnyd* 'that you know'), the second a purposive conjunction (*med od töždö* 'in order you don't sorrow'). (44a) displays *medby* 'in order to' with an infinitive adjunct, and *med oz* 'in order not to' with a finite predicate; (44b) presents an example of the contracted conjunction (*moz* < *med oz*), and (44c) is an example of the Komi-Permyak *štoby* in purposive function.

(43) Old Zyrian
Ön *tyr-je* *vôk-ö-jas* *gažal-am* [***med*** *töd-adnyd* *uź-amö-jas*
now time-ILL brother-1SG-PL want-PRS.1SG **COMP:OPT** know-PRS.2PL sleep-PP-PL
dor-yś], ***med*** *o-d* *töžd-ö,* *kučömkö* *i* *muköd* *vôjtyr,*
about-ELA **in.order** not.PRS-2 sorrow-CN.PL any also other people
kôd-jas-lön *abul* *kêžaśan.*
who-PL-GEN is.not hope
'Now, brethren, I want you to know about passed away (people) that you don't sorrow as others which have no hope.'
(L52: 65, lines 10–15 [1 Thessalonians 4, 13])

(44) a. KoZ
Seki *sijö* *nyl-ys-köd* *mun-am* *voj* *puk-an* *iń-śań* *uź-ny,*
then this girl-3SG-COM go-FUT.1PL night sit-PRP place-EGR sleep-INF
medby *tödmaś-ny* *i* ***med*** ***o-z*** *jöz-köd* *guľajt.*
in.order get.to.know-INF also **in.order** **not.PRS-3** people-COM walk.CN
'Then, from the evening sit-in, we go to sleep with this girl, in order to get to know each other and that she won't hang out with others.'
(Upper Vychegda; UK4: 148.2)

b. KoZ
Da *ybös-sö* *śipt-is,* ***moz*** *ńekod* *aďďžy.*
and door-ACC.3SG close-PST.3SG **in.order.not** nobody see.CN
'And he closed the door in order that nobody would see him.'
(Udora; UK3: 314.22)

c. KoP

No	*etö*	*źaḱ-lö*	*kol-ö*	*końćit-ny,*	***štöby***
but	this	son.in.law-DAT	be.necessary-PRS.3SG	finish-INF	**in.order**
sija	*imeńńö*	*e-z*	*laǵejt.*		
(s)he	property	not.PST-3	inherit.CN		

'But this son-in-law has to be finished in order that he shall not inherit the property.'
(Northern Permyak; UK1: 24.1)

Šutov (1999: 39) argues that in Udmurt gospels translated from Church Slavonic or Russian originals in 1847, the purposive conjunction which renders Russian *čtoby* is often not *šuysa*, but *med* 'let it be; shall be', or *medaz* 'let not be; shall not be' which, according to him, is used "in the role of a particle and conjunction", equivalent to Komi *med(ym)* (Šutov 1999: 39). He illustrates this with (45). In (45a) *med* is in clause initial position and behaves as in Komi. However, in (45b), the verb-final word order has the particle following the direct object (*tiľedyz* 'yee.ACC'), i.e. it is not in clause initial position. Be it as it may, this optative particle is not used to express purposive clause meaning in the modern Udmurt grammar, where, as we already know, *šuysa* is employed in this function. One may conclude that the development of the optative particle into a conjunction in Udmurt was initiated by translational works in the 19th century, but it did not gain ground and is not grammatical in modern Udmurt. For dialectal varieties, however, this development is reported, at least for purposive clauses (see Section 3.1 for an example).

(45) Ud.

a. *Oźy-en teľmyr-e aran kuźo-ly,* [***med*** *leź-oz*
so-INS pray-IMP.2PL harvest lord-DAT **in.order/COMP:OPT** send-FUT.3SG
araś-os-yz aran-vyl-az].
mower-PL-ACC harvest-on-ILL.3SG
'So pray to the lord of harvest that he will sent mowers to the harvest.'

b. *En kurlal-e, tiľedyz* ***medaz*** *kurlal-e.*
NEG.IMP2 judge-CN.PL yee-ACC **in.order.NEG** judge-CN.PL
'Don't judge in order not be judged yourself.'

(19th century, Šutov 1999: 39)

2.1.7 An epistemic complementizer

The Russian particle and complementizer *budto(by)* (see Hansen, Letuchiy & Błaszczyk, this volume) was borrowed into Komi as *byťťö(kö)*, and into Udmurt as *budto(ke)*, where the element *kö/ke* is the genuine conditional particle (see Section 2.2.3 below). The question to what degree this particle can be classified as a complementizer in Permic is tightly connected to the same question in Russian,

as both combine freely with the respective fact type complementizer. Unfortunately, this question cannot be answered in the present article.[4] For Komi, it will be shown in Section 3.2 that the particle combines freely with the fact-type complementizer *myj*. The basic similative meaning of *bytťö-kö* 'as if' is illustrated by (46a); this similative meaning is also present when the particle appears with a complement of the verb 'seem', as in (46b). With an acquisition of knowledge CTP, namely 'hear' in (47), the particle adds an epistemic (indirective) meaning. Example (47) contains also the quotative particle *pö*, but, as the minimal pair in (48) should demonstrate, the epistemic-indirective meaning of (47) is not due to the quotative particle. On the base of the Komi examples (46)–(48) one might assume that the element *kö* appears with the particle in its similative function, but not in its indirective marking function. However, the available evidence is insufficient to warrant such conclusion.

(46) a. KoZ
Da *menym* *gort-ö* *mun-ny* ***bytťökö*** *janďžim.*
and I.DAT home-ILL go-INF **as.if** shame
'And going home feels like shame to me.'
(Upper Vychegda; UK4: 192.31)

b. KoZ
Menym *kažiťć-ö* [***bytťökö*** *mam-ö* *menśim* *töd-ö*
I.DAT seem-PRS.3SG **as.if** mom-1SG I.ABL know-PRS.3SG
kupajťć-öm-ös.
bath-PP-ACC
'To me it seems as if mother knew about my swimming.'
(Upper Vychegda; UK4: 196.28)

(47) KoP
Marko *Bogatöj-lö* *kylś-as,* [***byďťö*** *jen-ys* *voll-öm*
Marko Rich-DAT be.heard-FUT.3SG **COMP.INDIR** god-3SG be.somewhere-PRF
pö]
QUOT
'Rich Marko learns that God had (allegedly) been around'
(Northern Permyak; R78: 414.19)

(48) KoZ (elicited)
a. *Marko* *viśtal-is,* [***myj*** *jen* *vol-öma* ***pö***].
Marko say-PST.3SG **COMP** God be-PRF.3SG **QUOT**
'Marko said that God was around.'

b. *Marko* *viśtalis,* [***bytťö*** *jen* *vol-öma* ***pö***].
Marko say.PST.3SG **COMP.INDIR** God be-PRF.3SG **QUOT**
'Marko said that, allegedly, God was around.'

4 It was not possible to study it properly because it hardly occurs in the studied texts, and the consultants used it either reluctantly (Komi), or not at all (Udmurt).

The epistemic-indirective meaning is also attested in (substandard) Udmurt; see *buttoke* in (49). However, the genuine Udmurt similative construction ('seems as if') does not make use of *budto(ke)*, but consists of the multifunctional auxiliary *poty-* 'come out' and the postposition *kaď* 'like'; cf. (50).

(49) Ud.
Vera *eš-ed-ly,* [*mon* ***butto-ke*** *viśə-śko*].
say.IMP.2SG friend-2SG-DAT I **COMP.INDIR** be.sick-PRS.1SG
'Tell your friend I'm (allegedly) sick.'
(Middle Čepca; Karpova 1997: 171)

(50) Ud. (elicited)
Boris *araky-ez* *u-g* *jaraty* ***kaď*** ***pot-e***.
Boris vodka-ACC NEG.PRS-3 love.CN **like** **come.out-PRS.3SG**
'It seems that (as if) Boris dislikes vodka.'

Due to its semantics, the Komi indirective particle/complementizer *byťťö(kö)* is prone to occur with the evidential 2nd past tense (cf. Cypanov 2005: 164). In principal, for Permic languages, one may well expect a genuine epistemic strategy which does not rely on a borrowed particle. This strategy would consist in the employment of perfect forms in the complement – the perfect (second past), and further past tenses built with the evidential/inferential particles, Komi *vylöm* and Udmurt *vylem* (see Cypanov 2005: 159–176; Leinonen 2000; Siegl 2004; Winkler 2011: 104). Such case is illustrated in (51) where the perfect tense on the predicate in a fact type *myj*-complement clause of the CTP 'know' may be read as carrying an inferential meaning (Siegl 2004: 109–114). A systematic account of the interaction of an epistemic conjunction based on the similative 'as if' and inferential and evidential readings of Permic tenses is a desideratum for future research.

(51) KoZ
A *mam-ö* *džik-pyr* *i* *gögörvo-is,* [***myj*** *žugöd-**öma***
but mom-1SG at.once also understand-PST.3SG **COMP** break-**PRF**
sakarńića].
sugar.holder
'But mom understood immediately that he had broken the sugar holder.'
(Upper Vychegda; UK4: 190.10, Siegl 2004: 113)

2.2 Other finite complements

2.2.1 Paratactic complements (and the question of complementizer omission)

Parataxis (or juxtaposition) means here that the complement-taking predication and the complemented predication are posed one after another without any subordination marker. Within this sequence, the prosodically disrupted second predicate is to be interpreted as a complement of the first (cf. SKJa. II: 250–251 for the intonational pause). Given that conjunctions are often borrowed in Permic, as well as generally in Uralic languages – cf. the overview by Leinonen (2002) on the borrowed Russian conjunctions in Uralic languages originally collected by Maitinskaja (1982, chapter 2) – Uralists consider paratactic structures as an older strategy employed before borrowing of conjunctions set in (Tauli 1966: 87f.; Honti 1994: 95; Šutov 1999: 42; Osnovy 1974: 215). In relation to the other complementation strategies, one may either interpret parataxis (according to the Uralist perspective) as a basic structure, from which by grammaticalization of WH-words, particles and other structures with formal complementizer elements developed, or one may interpret paratactic complements as in (2a) and (2b), repeated here as (52a) and (52b), as cases of complementizer omission. The pause between the matrix clause and the complement clause is orthographically rendered with a comma (as in the above examples), or with a colon, not only with utterance verbs, but also with perceptions verbs, as in (53) (cf. also SKJa. II: 250).

(52) a. KoZ
Ćajt-a, [*skörm-as*].
think-PRS.1SG be.angry-FUT3SG
'I think she'll be angry.'
(R96: 114.5)

b. Ud.
E, *ćakla-śko,* [*jyr-e* *byr-iz*].
INT think-PRS.1SG head-1SG get.lost-PST.3SG
'Hey, I think I lost my head.'
(OU1: 38.6)

(53) KoZ
Ańuta *pyr-is* *o-š* *kerka-ö.* *Da* *aďž-ö:* [*seni*
Anyuta enter-PST.3SG bear house-ILL and see-PRS.3SG there
pukal-ö *da* *börd-ö* *as-las* *ićet* *ćoj-is* *Mašö*].
sit-PRS.3SG and cry-PRS.3SG self-GEN.3SG little sister-3SG Masha
'Anyuta entered the house of the bear, and she sees there sits her little sister Masha and cries.'
(Upper Vychegda; UK4: 424.18)

A paratactic object complement clause may well appear in a SOV structure, as in (54) from Udmurt. It was discussed above whether word order is a criterion for the obligatory use (resp. omissibility; see Section 2.1.1 above) of the Udmurt complementizer *šuysa*. As Vilkuna (1998: 185) expressed it, the complementizer is obligatory when the object sentence complement takes the O position in an SOV structure, as in (21a) repeated here as (55a), whereas it is facultative when the complement clause follows the matrix sentence. This is a tendency, not a grammatical rule, as the counterexample (55b) with *šuysa* in SOV structure reveals.

(54) Ud.
Slavjan-jos [*každoj gurt-yn gurt kuźo vań*] *koža-lo*
Slavonian-PL every house-INE house ghost there.is think-PRS.3PL
vyl-em.
be-PRF.3SG
'The Slavonians believed that there is a domestic ghost in every house.'
(Winkler 2011: 163)

(55) Ud.
a. *Nina* [*mon ćoryg śi-iśko*] ***šuysa*** *malpa-z*.
Nina I fish eat-PRS.1SG **COMP** think-PST.3SG
'Nina thought that I eat fish.'
(Vilkuna 1998: 185)
b. *Epifańij* [*vań-ze no aďďži-śko*] *šu-em*
Epifanij all-ACC.3SG but see-PRS.1SG say-PRF.3SG
'Epifanij said he sees everything.'
(Winkler 2011: 170)

It is a matter of future research to determine the more exact parameters of zero complementizers in Permic. One of the parameters which must be paid attention to is the functional capacity of a complementizer. The fact type complementizer 'what' is, obviously, capable of marking a proposition as a complement without a CTP. With an elliptic utterance predicate as in (56) this function comes close to the function of a quotative particle. What exactly then would be omitted if (56) would come without *što*?

(56) KoZ
No sija ötkažićć-is, [***što*** *"me vörav-ny o-g mun"*].
but (s)he stand.back-PST.3SG **COMP** I hunt-INF NEG.PRS-1 go.CN
'But he stood back (saying) "I won't go hunting".'
(Upper Vychegda; UK IV: 150.21)

2.2.2 Polar questions complements

Polar question (yes-no question) complements make use of two independent, but combinable strategies. One strategy consists in the application of an enclitic question particle, Komi *-ö*, Udmurt *-a*, on the complement predicate as in (57), or, in principle, on any other constituent of the complemented clause. The other strategy is a negative alternative structure. In this structure, the complement predicate is immediately followed by a respective form (tense, person) of the negation verb. Between the two verb forms there is an intonational pause, or an explicit indicator of alternative 'or', cf. (58a) and (58b). The two patterns may be combined as in (58c). The variants in (58b) and (58c) are considered synonyms. Instead of finite verbs, the existential and the negative predicative particle may appear as alternatives; see (59).

(57) Question particle
a. KoZ (elicited)
Anna *jual-is,* [*giž-is-**ö*** *zonm-ys* *piśmö-sö*].
Anna ask-PST.3SG write.PST.3SG-**Q** boy-3SG letter-ACC.3SG
b. Ud. (elicited)
Anna *jua-z,* [*pijaš* *gožtet-se* *gožtiz-**a*** *[šuysa]*].
Anna ask-PST.3SG boy-3SG letter-ACC.3SG write.PST.3SG-**Q** COMP
'Anna asked whether the boy wrote the letter.'

(58) Negative alternative
a. KoZ (elicited)
Me *vunöd-i,* [***lokt-as*** ***o-z*** *Anna* *aski*].
I forget-PST.1SG **come-FUT.3SG** **NEG.PRS-3SG** Anna tomorrow
b. Ud. (elicited)
Mon *vunet-i,* [*Anna* *čukaźe* ***lykt-oz*** ***jake*** ***u-z***
I forget-PST.1SG Anna tomorrow **come-FUT.3SG** **or** **NEG.PRS-3SG**
[šuysa]].
COMP
c. Ud. (combined strategy; elicited)
Mon *vunet-i,* [*Anna* *čukaźe* ***lykt-oz-a,*** ***u-z-a***
I forget-PST.1SG Anna tomorrow **come-FUT.3SG-Q** **NEG.PRS-3SG-Q**
[šuysa]].
COMP
'I forgot whether Anna will come tomorrow or not.'

(59) KoZ
Petal-ö, *vidl-ö,* [***em*** ***abu***]*!*
go.outside-IMP.2PL look-IMP.2PL **there.is** **there.is.not**
'Go outside and look whether he's there or not!'
(Upper Vychegda; UK4: 152.16)

Instead of the genuine question particles one may find the borrowed Russian question particle *-li*, as in (60). Note that this is an example for heavy Russian influence on Komi since not only is the question particle Russian, but also the complement taking adjective predicate '(be) known, obvious' (Russian *izvestno*), as well as the complement predicate 'succeed' (Russian *udatsja*). In Udmurt, other, non enclitic, question particles (conjunctions) are used in the same function, e.g. *meda* 'maybe, whether' in (61a) which may be combined with the question particle *-a* occurring on a constituent of the complement clause, as in (61b); *meda* itself cannot take the question particle *-a*.

(60) KoZ

Korśl-am	*da*	*iźvesvö*	*lov-ö,*	[*udajćć-is-**lĭ***	*e-z-**lĭ***].
search-FUT.1PL	and	known	become-PRS.3SG	succeed-PST.3SG-**Q**	NEG.PST-3-**Q**

'We'll search and it will turn obvious whether it succeeded or not.'
(Upper Vychegda; UK4: 410.5)

(61) a. Ud.

Ku-ze	*Sarapul-e*	*leďžga-zy,*	[*puny-len*	*övöl*	***meda***	*šusa*].
skin-ACC3SG	S.-ILL	send.PST-3PL	dog-GEN	isn't	**Q**	COMP

'They sent the (wolf) skin to Sarapul, (in order to find out) whether it's not a dog's one.'
(OU2: 226.6)

b.

Ku-ze	*Sarapul-e*	*leďžga-zy,*	[*puny-len*	*övöl-**a***	***meda***	*šusa*].
skin-ACC.3SG	Sarapul-ILL	send.PST-3PL	dog-GEN	isn't-**Q**	**Q**	COMP

'They sent the (wolf) skin to Sarapul, (in order to find out) whether it's not a dog's one.' (elicited)

2.2.3 Conditional complements

When Permic conditional complements are characterized by a special complementizer, it is the borrowed Russian conjunction *esli* 'if' as in (62a)–(62b). But note that the Komi example (62a) has the conditional word *jeśĺĭ* following the general complementizer *myj*, and in the Udmurt example (62b) we find an instance of a pleonastic, so called "framing" construction in which the borrowed conditional conjunction *jesĺĭ* appears at the beginning of the complement and the genuine conditional particle *ke* follows the verbal predicate. Differently, in examples (10a) and (10b) from the introduction there are no borrowed conditional markers. Based on these examples one must state that a conditional sentence in Permic does not take special form when it comes as a complement. The conditional particle *kö/ke* is enclitic and not prone to move to the clause initial position where it could take the function of a complementizer. Still, it can appear with the first constituent of a complement clause, as in (63).

(62) a. KoP
Pöriś-žyk *soj-ys* *viśtal-is* *Ivan-lö,* [***myj*** ***jeśl̀i*** *sijös,* *Ivan-ös*
old-CMP sister-3SG tell-PST.3SG Ivan-DAT **COMP** **if** (s)he.ACC Ivan-ACC
o-z *lept-ö* (...)]
NEG.PRS-3 lift-CN.PL
'The eldest sister told Ivan that if they refused to lift him up (...).'
(Northern Permyak; R78: 428.7)

b. Ud.
Nu *malpa-śkom,* [***jesl̀i*** *puny-zy* *lu-is* ***ke,*** *fśo,* *byr-im*].
well think-PRS.1PL **if** dog-3PL be-PST.3SG **if** all perish-PST.1PL
'Well, we think, if they have a dog, it's the end, we are finished.'
(OU2: 45.3)

(63) KoZ
Važön *pöriś* *l̀ud-ys* *šulyvl-ömaöś,* [*sijö* ***kö*** *pukt-an* *da*
ago old people-3SG use.to.say-PRF.3PL (s)he.ACC **if** put-PRS.2SG and
krańit-an, *o-z* *lo* *sy* *vo-ö* *ńekučöm* *viśöm*].
keep-2SG NEG.PRS-3 be.CN this year-ILL no kind desease
'The old time people used to say that if you put it (dried mink over the window) and keep it there won't be any kind of desease in this year.'
(Udora; UK3: 326.2)

Example (64) displays expressions of irreality. In (64b) the Udmurt variant has the matrix verb in the conditional mood (in which the distinction of person is neutralized) and the complement clause is not specified for irrealis. The Komi variant in (64a), on the other hand, has an irrealis particle in the matrix clause, and the complement clause contains the conditional particle *kö*.

(64) a. KoZ (elicited)
Me ***eśkö*** *töd-i,* [*völ-i* ***kö*** *Garri Potter-ys* *menam* *kerka-yn*].
I **IRR** know-PST.1SG be-PST.3SG **if** Harry Potter-3SG I-GEN house-INE

b. Ud. (elicited)
Mon *tody-**sal***, [*Garri Potter* *mynam* *korka-n* *val* *šuysa*].
I know-**COND** Harry Potter I-GEN house-INE has.been COMP
'I would know if Harry Potter would have been at my house.'

2.3 Non-finite complements

Non-finite complements have restricted syntactic possibilities concerning the expression of complex propositions (arguments, circumstantials), although Udmurt participle clauses are more flexible than those in Komi. The most restricted non-finite verb form is the infinitive (Section 2.3.1). Participles are a frequent and diverse subordination instrument (Section 2.3.2), and converb complements are more important in Udmurt than in Komi. Generally, due to the

influence of Russian, subordination of the agglutinative type, involving converbs and participles, has become less important in Komi than in Udmurt (cf. Serebrennikov 1967, Leinonen 1998).

2.3.1 Infinitive complements

The Permic infinitive is a verbal substantive, an action noun which may function as an argument, adjunct or complement. It is also used as the lexical citation form of verbs. The infinitive ending is *-ny* in all Permic languages. There is no overt expression of subject argument on the infinitive, which is coreferential with the subject (or with the object in a causative construction). Subject indicating possessive suffixing of the infinitive as in Hungarian – e.g. *men-ne-m kell* 'I have to go (go-INF-1SG is.necessary)' – is reported for Komi; e.g. *mun-ny-m og vermy* 'I can't go (*mun-ny-m* go-INF-1SG)' by Rédei (1978: 112). However, these forms are random compared to the pragmatically motivated possessive suffix marking in Komi; cf. the 2nd person singular possessive suffix *-(y)d* in (66) and (67), or the possessive suffixes of the 3rd person singular (nominative *-(y)s*, accusative *-sö*) in (68a) and (68b).[5] As a final adjunct the infinitive occurs with any kinds of verbs (e.g. verbs of movement). In both Permic languages, the infinitive is a complement of knowledge predicates. In (65) the infinitive functions as a subject complement of the Komi CTP *vun-* 'become forgotten' with a genitive experiencer. With the transitive CTP *vunöd-* 'forget' in (66) the infinitive functions as an object complement. Otherwise as in some existing analysises of "nominal object sentences" (Stipa 1960), the infinitive complement should not be analyzed as a direct object, because it is not marked as an object: (66)–(67) have the possessive suffix in the nominative, but possessive direct objects are obligatorily object marked in Permic (cf. Klumpp

5 The pragmatic function of the 3rd person singular possessive suffix can be illustrated more precisely by (ia) and (ib) below, where the suffix appears with independent infinitives. In (ia) we find the accusative variant of the possessive suffix as a focus marker, whereas in (ib) we find the nominative variant of the suffix as a marker of givenness.

(i) Komi
a. *Seśa pudjas-öny,* [*kyčöm parťija-ly kyjal-ny-**sö***]$_{\text{Foc}}$.
then allot-PRS.3PL which party-DAT catch-INF-**ACC.3SG**
'Then they allot [which party has to CATCH]$_{\text{Foc}}$.'
b. *Kod-ly uś-ö kijal-ny-**s**, ńija vötlyś-öny meć bör-śys.*
who-DAT fall-PRS.3SG catch-INF-**NOM.3SG** they run-PRS.3PL ball after-ELA.3SG
'Those who have to catch (lit.: to whom the [aforementioned] catching falls), they run after the ball.'
(MS; UK3: 56)

2009, 2014). Unlike participles which, in order to depend on a verb, take one of its argument positions, the infinitive is already able to depend on another verb. Infinitive complements are thus natural complements. In both Permic languages the infinitive is also a complement of a range of modal and phasal CTPs, as e.g. in (68) and (69).

(65) Ko.
Me *zē* *ziľa* *ćeťć-i,* *da* *veśig* *menam* *vun-öma* [*myśśy-ny*].
I very quickly get.up-PST.1SG and nearly I.GEN get.lost-PRF wash-INF
'I got up very quickly, and even forgot to wash.'
(UV; UK4: 206.26)

(66) KoZ
A [*pyvśy-ny-d*] *taj* *vunöd-innyd?*
and bath-INF-NOM.2SG MOD forget-PST.2PL
'And bathing, you seem to have forgotten about?'
(Bubrix 1949: 54)

(67) KoZ
dugdy [*vina-tö* *juny-d*]
stop vodka-ACC.2SG drink.INF-NOM.2SG
'stop drinking vodka!'
(Stipa 1960: 172 < WU 374)

(68) KoZ
a. *No* *götyr-ys* *ńinöm* *e-z* *vermy* *viśtal-ny-s*
but wife-3SG nothing NEG.PST-3 can.CN say-INF-NOM.3SG
'But the wife couldn't say anything.'
b. *No* *sijö* *ńinöm* *e-z* *vermy* *viśtal-ny-sö*
but (s)he nothing NEG.PST-3 can.CN say-INF-ACC.3SG
'But he couldn't say anything.'
(Middle-Sysola; UK3: 40–42)

(69) KoZ (elicited)
a. *Zonm-ys* *verm-as /* *zavoďit-is* [*giž-ny* *piśmö-sö*].
boy-3SG can-FUT.3SG begin-PST.3SG write-INF letter-ACC3SG
b. Ud. (elicited)
Pijaš [*gožtet-se* *gožty-ny*] *bygat-e /* *kutsk-iz.*
boy letter-ACC.3SG write-INF can-PRS.3SG begin-PST.3SG
'The boy is able / started to write the letter.'

As a complement of propositional attitude predicates (e.g. 'think'), the infinitive stands in opposition to fact type complements: it triggers purposive reading of the verb ('intend'), cf. (70a) and (70b) for Komi *dumajt-* 'think', and (71a) and (71b) for Udmurt *malpa-* 'think'.

(70) Komi *dumajt-* 'intend ~ think'
a. *Pop* *dumajt-is*, [*Jemöl'-ös* *vi-ni*].
priest think-PST.3SG Yemel-ACC kill-INF
'The priest thought (intended) to kill Yemel.'
(UV; UK4: 6.30)
b. *Najö* *dumajt-öny*, [*mun-am* *kö* *ötlayn*, *siďžkö* *žöńik* *da* *ńevesta*].
they think-PRS.3PL go-1PL IRR together so.as.if groom and bride
'They think we are together like groom and bride.'
(PM 1112)

(71) Ud. *malpa-* 'intend ~ think'
a. *Ton* [*mam-me* *vi-ny*] *malpa-d-a*, *mar-a?*
you mom-ACC.1SG kill-INF think-FUT.2SG-Q what-Q
'Do you intend to kill my mother, or what?'
(OU1: 46.18)
b. [*Mašina-my* *övöl* *ńi* *šusa*] *malpa-škom* *val.*
car-1PL isn't already COMP think-PRS.1PL was
'We've been thinking that we had no car anymore.'
(OU2: 45.8)

Finally, the infinitive also appears with CTPs like 'see', as in (72); here the complement contains an independent infinitive. The modal function of the independent infinitive (with a subject in dative case) is also well known in Russian, and may be interpreted as Russian structural influence in Permic.[6]

(72) KoZ
Me *vör-ö* *voš-i* *da* [*gort-ö* *mun-ny*] *o-g* *aďďži.*
I forest-ILL get.lost-PST.1SG and home-ILL go-INF not.PRS-1 see.CN
'I got lost in the forest and don't find (how/where) I should go home.'
(Upper Vychegda; UK4: 424.14)

2.3.2 Participial complements

Participles are verbal adjectives which are used in argument as well as in attributive function. There are two participles to be considered here: the perfect participle in *-m* which is used in different case forms, and the present participle in *-n*

6 A non embedded occurrence of the independent infinitive would be:

(i) KoP
Siśśa *myj* *ḱerny* *menym!*
then what do.INF I.DAT
'What should I do then (lit. Then what to do to me)?'
(Northern Permyak; UK1: 34.13)

with a smaller range of uses. (For another participle, namely the agent participle in *-yś*, see the peripheral complementation strategies in Section 2.6). In Udmurt, participial complements may contain the same dependent pronouns and adverbs as a finite clause (e.g. Šutov 1999: 25); for the participle in Komi see Cypanov (1997), and for a comparison of the Permic and Mari participles see Brykina and Aralova (2012). The perfect participle ends in *-m* in all Permic languages (Ko. *-öm*, Ud. *-em*, *-m*). It occurs in argument function and can thus serve as a subject or object sentence complement. From the point of view of complementation, the perfect participle, is the most important and frequent participle; see the examples in (73). The subject/agent can be expressed overtly as a possessor in the genitive or ablative case.[7] A participial complement which functions as direct object of a CTP is inflected according to the rules of Permic differential object marking, i.e. the participial complement is either an unmarked object, or marked by the accusative variant of a possessive suffix (cf. Klumpp 2009, 2014). Note that GSUJa. (1970: 165) treats participle (and also infinitive) complements in the volume on the simple sentence as direct objects. Usually, but not obligatorily, the agent-possessor is cross-referenced by a possessive suffix on the participial complement. While in Udmurt the possessive suffix expresses coreferentiality, as in (73b) and (74a), in Komi its use is in most cases pragmatic, see (73a) and (74b). The participle in (74b) may either be interpreted as a complement predicate or as a lexicalized *m*-form. More clear instances of the latter are the plural marked participle subject in (75a), or (75b) where the agent is not expressed as a possessor, but as a determinative first part of a compound.

(73) a. KoZ (elicited)

Anna	*aďďž-ö*	[*na-lyś*	*pińaś-öm-sö*]	*da*	*śeral-ö.*
Anna	see-PRS.3SG	they-ABL	quarrel-PP-ACC.3SG	and	laugh-3SG

b. Ud. (elicited)

Anna	[*soos-leś*	*tyškaśk-em-zes*]	*aďďž-e*	*no*	*śerekja.*
Anna	they-ABL	quarrel-PP-ACC.3PL	see-PRS.3SG	and	laugh.3SG

'Anna sees them quarrel and laughs.'

(74) a. Ud.

Kwaź	*šu-em:*	*"Aďďž-i*	*mon*	[*kytsy*	*pon-em-de*]*!"*
god	say-PRF.3SG	see-PST.1SG	I	where	put-PP-ACC.2SG

'God said: "I saw where you put it!"'
(Wichmann 1901: 125)

7 In all Permic languages, the choice between genitive and ablative depends on the syntactic function of the possessum: if it is a direct object, the case of the possessor switches from genitive to ablative.

b. KoZ
Vaśiľej-ly *vöjćća* *kyvźiś-öny* [*vo-gögör-sa* *ol-öm-tö*].
V.-DAT upon listen-PRS.3PL year-round-ADJ live-PP-ACC.2AG
'Upon Saint Basil's day they listen for the life of the whole year.'
(Udora; UK3: 332.9)

(75) Komi lexicalized participles
a. KoZ
Vödžö *kyvlyvl-ö* *ströjćć-öm-jez*.
further be.audible-PRS.3SG build-PP-PL
'Further can be heard construction works.'
(Udora; UK3: 332.15)
b. KoZ
ji *kilal-öm*
ice drift-PP
'ice drift'
(Upper Vychegda; UK4: 192.2)

Further constructions with the *m*-participle have it in specific case forms (dative, elative, ablative, instrumental) or with different postpositions. These forms, sometimes understood as complex converbs, are the following: (i) the dative-marked participle as a complement of the verb 'believe' in Udmurt, as in (76); (ii) the elative-marked participle, Komi *-öm-yś*, Udmurt *-e/am-yś* functions as a complement of phasal verbs which mean 'end, stop', see (77a) and (77b); (iii) In Komi, the elative marked participle is also used as a complement of the verb 'fear', whereas in Udmurt we find the ablative here, cf. (78a) and (78b). In both languages the agent of the complement predicate is formally a genitive possessor. In Komi, the elative participle has also the causal reading 'out of doing x', and may appear as a circumstantial adjunct with any verb. (iv) The instrumental-marked participle, Komi *-öm-ön*, Udmurt *-em-en*, has an essive meaning, and appears as a complement of utterance verbs in Komi, see (79a); it is also used in circumstantial expressions in its basic instrumental meaning, see (79b) (this function is called "Modus obliquus" by Stipa 1960: 196). (v) In Udmurt, the *m*-participle plus the possessive suffix appears in a desiderative construction, as in (80), and (iv), without the possessive suffix, in a construction with the auxiliary *kary-* 'make, do' meaning 'pretend', see (81).

(76) Ud. (elicited)
Mari *osk-e* [*Xristos-len* *beren* *vu-em-ez-ly*].
Mari believe-PRS.3SG Christ-GEN back come-PP-3SG-DAT
'Mary believes in the returning of Christ.'

(77) a. KoZ
I *dugd-as* [*śin-tö* *kuröd-öm-iś*].
and stop-FUT3SG eye-ACC2SG burn-PP-ELA
'And it will stop burning in the eyes.'
(Upper Vychegda; UK4: 96.15)
b. Ud. (elicited)
Pijaš [*gožtet-se* *gožt-em-yś*] *dugd-iz.*
boy letter-ACC.3SG write-PP-ELA stop-PST.3SG
'The boy stopped writing the letter.'

(78) a. KoZ (elicited)
Me *pol-a* [*sy-lön* *lokt-öm-śy-s*]
I fear-PRS.1SG (s)he-GEN come-PP-ELA-3SG
b. Ud. (elicited)
Mon *kyška-śko* [*so-len* *lykt-em-ez-leś*].
I fear-PRS.1SG (s)he-GEN come-PP-3SG-ABL
'I'm afraid of his coming.'

(79) a. KoZ
Śemuk Śem-ös *šu-i-sny* [*vij-öm-ön* *Puzla Öńö-jas-ös*]
Syemuk Syem-ACC say-PST-3PL kill-PP-INS Puzla Önyö-PL-ACC
'They said that Syemuk-Syem had killed Puzla-Önyö and his fellows.'
(Upper Vychegda; UK4: 76.6)
b. KoZ
Seki *oš-sö* *kyč* *pij-ö* *bośt-öni* *lyž-ön* *kyčot-öm-ön.*
then bear-ACC.3SG circle in-ILL take-PRS.3PL ski-INS circuit-PP-INS
'Then they surround the bear by circuiting on ski.'
(Upper Vychegda; UK4: 84.27)

(80) Ud. (elicited)
Pijaš-len [*gožtet-se* *gožt-em-ez*] *pot-e.*
boy-GEN letter-ACC.3SG write-PP-3SG come.out-PRS.3SG
'The boy wants to write the letter.'

(81) Ud. (elicited)
Pijaš [*gožtet-se* *gožt-em*] *kar-e.*
boy letter-ACC3SG write-PP make-PRS3SG
'The boy pretends to write the letter.'

Among the forms consisting of the participle and postposition, one should mention the combinations with the Komi postposition *jyl-yś* ~ *jyvśys* 'about'; the postposition comes in elative case (*jyl-yś* 'over-ELA' ~ *jyv-ś-ys* 'over-ELA-3SG). Postpositional 'about' formations appear as complements of verbs of knowing and telling, as in (82) and (83). Other formations with postpositions are temporal (*-öm böryn* 'after x-ing'), causal (*-öm vösna* 'because of x-ing'), or purposive (*-öm vylö* 'in order to x').

(82) KoZ (elicited)
Vunöd-i [*sy-lön* *lokt-öm* ***jyvśys***], *da* *e-g* *vöv* *gort-yn.*
forget-PST.1SG (s)he-GEN come-PP **about** and not.PST-1SG be.CN home-INE
'I forgot about her coming and was not at home.'

(83) KoZ
Sy-ly *viśtal-isny* [*kazna* *voš-öm* ***jylyś***].
(s)he-DAT tell-PST.3PL cashier get.lost-PP **about**
'They informed him about the loss of the cash box.'
(FF: 1219)

The present participle has more restricted distribution. In Udmurt, the present participle in *-on* is used in the same function as the perfect participle, if the situation requires it. The equivalent form in Komi, *-an*, is not used for this purpose – instead the anterior temporal value of the *m*-participle is ignored (or a bifinite complement strategy with the standard complementizer *myj* is used); cf. (84a) and (84b). Serdobol'skaja et al. (2012: 413) report for the Beserman dialect of Udmurt that the participle in *-on* is used only with phasal verbs; this means that in variants of Udmurt the use of this participle as a complement is also restricted. The functional opposition between participle and finite complements will be examined in Section 4.

(84) a. Ud. (elicited)
Mon [*Anna-len* *asky* *lykt-on-ez*] *śaryś* *vunet-i.*
I Anna-GEN tomorrow come-PRP-3SG about forget-PST.1SG
b. KoZ (elicited)
Me *vunöd-i* [*Anna-lön* *aski* *lokt-öm*] *jyv-śy-s.*
I forget-PST.1SG Anna-GEN tomorrow come-PP about-ELA-3SG
'I forgot about Anna's coming tomorrow.'

2.3.3 Predicative particle complements

Both Permic languages have predicative particles; examples include the existential predicates, Udmurt *vań*, Komi *em* 'there is, there exists', or their negative counterparts, Udmurt *övöl*, Komi *abu*. Howbeit, only in Udmurt these particles can function as complements with the inflectional possibilities available for participles; cf. (85) and (86). The same structure is also found with predicatively used adverbials, as in (87).

(85) Ud. (elicited)
Anna [*gožtet-leś* *vań-ze*] *aďďž-iz.*
Anna letter-ABL there.is-ACC.3SG see-PST.3SG
'Anna saw there was a letter.'

(86) Ud. (elicited)
Pijaš-leś [*gožet* *gožty-ny* *kul-e-ze*] *tod-iśko.*
boy-ABL letter write-INF be.necessary-PRS.3SG-ACC.3SG know-PRS.1SG
'I know that the boy has to write a letter.'

(87) Ud.
[*Tymet-leś* *art-yn-ze*] *mon* *ton-tek* *no* *tod-iśko.*
pond-ABL close-INE-ACC.3SG I you-ABE also know-PRS.1SG
'That the pond is nearby I also know without you.'
(Šutov 1999: 26)

2.3.4 Converb complements

Converbs are verbal adverbs. Converb syntax is rich and varied in Ural-Altaic languages. Permic languages have a range of converbs with different functions, but generally speaking, temporal and modal subordination and chaining by converbs (as in the Altaic languages) is more common in Udmurt than in Komi. There are four converbs to be mentioned in connection with complementation: (i) the Udmurt general converb, (ii) the Udmurt necessive converb (participle), (iii) the Permic abessive converb, and (iv) the Permic *connegative*. The Udmurt general converb in *-sa* functions as a complement with the CTP 'finish', as in (88a). The corresponding construction in Komi makes use of the infinitive, as in (88b). The Udmurt necessive participle in *-ono* appears predicatively in purposive clauses with the conjunction *šuysa*, as in (89a). It it is rather rare in complement position, but see (89b). The Komi abessive converb *-tög*, occurs as a complement with a rather marginal class of CTPs, namely verbs meaning 'remain, stay'. The Udmurt abessive converb *-tek*, however, has a wider distribution. In addition, it functions as a negative counterpart of the infinitive with a range of CTPs as, e.g. 'can' in (90b). And finally, the negative stem, or the connegative, as in the matrix sentence of (1) and many other examples, may be interpreted as a converb complement restricted exclusively to complements of the negation verb.

(88) a. Ud. (elicited)
Pijaš [*gožtet-se* *gožty-sa*] *bydt-iz.*
boy letter-ACC.3SG write-CVB finish-PST.3SG
b. KoZ (elicited)
Zonm-ys *pomal-is* [*giž-ny* *piśmö-sö*].
boy-3SG finish-PST.3SG write-INF letter-ACC.3SG
'The boy finished writing the letter.'

(89) a. Ud.
So *magaźin-e* *lykt-iz,* *ńań* *baśt-ono* *šuysa.*
(s)he shop-ILL come-PST.3SG bread buy-PNEC COMP
'He came into the shop in order to buy bread.'
(Winkler 2011: 116)

b. Ud.
Ĺida-je *kńiga* *baśt-e* *ke,* [*soje* *lyďď-ono*] *koža-śko.*
Lida-1SG book buy-PRS.3SG if it.ACC read-PNEC think-PRS.1SG
'Whenever my Lida buys a book I think I have to read it.'
(OU2: 217.12)

(90) a. Ud.
Oźy *ju-os* [*kiśma-tek* *ik* *busy-os-y*] *kyľyl-izy.*
so corn-PL ripen-CVB:ABE same field-PL-1SG stay-PST.3PL
'Thus, my fields remained with the corn unripening.'
(Hamari 2011: 74–75; http://old.udmdunne.ru/articles/art265.html)

b. Ud.
So *bəgat-e* [*ləktə-nə /* *ləktə-tek*].
(s)he can-PRS.3SG come-INF come-CVB:ABE
'He can come/cannot come.'
(Beserman; Serdobol'skaja et al. 398)

2.4 Argument raising

In subject to object raising, what is semantically a subject argument of the complemented clause is syntactically realized as an object argument of a CTP, i.e. the subject argument of the complement clause is at the same time object argument of the CTP. In both Komi and Udmurt, there are substandard formations like the ones shown in (91) and (92). Winkler (2011: 170) reports the raising structure as facultative – cf. (93) – , adding that the range of this phenomenon in Udmurt is unknown (cf. also Šutov 1999: 43–44; Stipa 1960: 182–184). A systematic account of this phenomenon must consider parallel structures in Tatar (cf. Serdobol'skaja & Toldova 2011: 290). With CTPs 'take as, consider somebody to be something', it seems to be employed rather systematically: cf. its occurrence in the Komi construction *šu-* 'say' + *-ömön* in (92) and in the Udmurt example in (94) with an adjective predicative. In the contemporary varieties, there seem to be a blending of strategies. Accordingly, in (95) from Komi, as an alternative to the structure with the elative case, a raised argument combines with the conjunction introduced complement.

(91) Ud.
Brat-jos-ys [*so-je* *śut-em* *ku-oz*] *mawp-ō* *va.*
brother-PL-3SG (s)he-ACC hunger-PP die-FUT.3SG think-3PL was
'The brothers had thought him to have died of hunger.'
(OU1: 207.13)

(92) KoZ
Śemuk Śem-ös *šu-isny* [*vij-öm-ön* *Puzla-Ӧńö-jas-ös*] *sy ponda, myj …*
Semuk-Seme-ACC say-PST.3PL kill-PP-INS Puzla-Önyö-PL-ACC because
'They said that Semuk-Seme had killed Puzla-Önyö and his fellows because …'
(Upper Vychegda; UK4: 76.6)

(93) Ud.
Mi [*tone/ton* *byr-id* *[šuysa]*] *kožamy.*
we you.ACC/you.NOM vanish-PST.2SG COMP think.PST.1PL
'We believed you to have vanished/We thought that you have vanished.'
(Winkler 2011: 170)

(94) Ud.
… finľandi-yn *18.* *daur-e* *no* *aj* [*finn* *kyl-ez*
Finnland-INE 18. century-ILL and but Finnish language-ACC
ivrit, *vaškala* *grek* *no* *got* *kyl-jos-yn* *og-vyžy-yś*]
Hebrew old Greek and Gothic language-PL-INS one-root-ADJ
koža-ĺĺ-a-zy.
think-FRQ-PST-3PL
'… but in Finnland, in the 18. century, Finnish was believed as of one root together with Hebrew, Old Greek, and Gothic.'
(Suomi 493)

(95) KoZ
Vańa *Peťa-ös /* *Peťa-lyś* *töd-e,* [*myj* *[sija]* *lokt-is*].
Vanja Petja-ACC Petja-ABL know-PRS.3SG COMP (s)he come-PST.3SG
'Vanja knows that Petja came.'
(Pečora; Serdobol'skaja et al. 2012: 436)

2.5 Quotative particle and reference strategy

The quotative particle, Komi *pö*, Udmurt *pe*, is inserted in direct speech to mark it as quoted speech. Stipa (1960: 194) contrasts quoted speech, which "needs" the quotative particle, with the perfect tense as a *modus obliquus* which does not "need" a quotative particle because it already includes the notion of indirectivity. The particle may be inserted once or several times per quoted unit. One may ask whether with utterance verbs the quotative particle is not a complementizer rather than an accompanying phenomenon. Since it freely combines with the comple-

mentizer *myj* in Komi – cf. (96) –, as well as with *šuysa* in Udmurt – cf. (97) –, and since it does not move to the clause initial position in Komi, it is not. In addition to *pö*, Komi also has the autoquotative particle *miśa* (see Section 3.2; cf. also Serdobol'skaja et al. 2012: 387, 459).

(96) KoZ (elicited)

a.	*Anna*	*šu-öma,*	(*myj*)	[*lokt-a*	***pö***].
	Anna	say-PRF.3SG	COMP	come-FUT.1SG	**QUOT**
b.	*Anna*	*šu-öma,*	(*myj*)	[*lokt-as*	***pö***].
	Anna	say-PRF.3SG	COMP	come-FUT.3SG	**QUOT**

'Anna has said that she will come.'

(97) Ud.

[*"Kin-leś-o*	*pegďž-it?"*	***pe***],	*šusa*	*jua*.
who-ABL-Q	run.away-PST.3SG	**QUOT**	COMP	ask.PST.3SG

'"Who did she run away of" he asked.'
(OU1: 118.18)

As can be observed in (96a) and (96b), Komi has two possible reference strategies in quoted speech: the deictic strategy in (96a) in which the person of the complement predicate corresponds to the *deictic origo* in direct speech, and the anaphoric strategy in (96b) in which the person of the complement predicate has been accommodated to the matrix subject. These alternatives are available also in Udmurt (e.g., Šutov 1999: 44), as can be seen in (98) from the Besermyan dialect. Serdobolskaja and Toldova (2011: 288) report that the deictic strategy is the most preferred pattern with utterance verbs as in (96) and (98), but less preferred with emotional and causation verbs (and totally excluded in purposive and causal clauses, which also employ the subordination marker *šuysa*). This holds for subjects; other constituents of dependent clauses (direct and indirect objects) are, according to their observations, only rarely encoded along the deictic strategy.

(98) Ud.

anaj-ez	*vera-z*	*ataj-ez-lə*	[*mone /*	*soje*	*kwaret-iz*	
mother-3SG	tell-PST.3SG	father-3SG-DAT	I.ACC	(s)he.ACC	scold-PST.3SG	
kužəj-e /	*kužəj-ez*]	*šuəsa*				
boss-1SG	boss-3SG	COMP				

'The mother$_i$ told the father that my$_i$/her$_i$ boss had scolded me$_i$/her$_i$.'
(Besermyan; after Serdobolskaja & Toldova 2011: 287)

2.6 Peripheral complementation strategies

The strategies mentioned in this section are not complementation strategies in the strict sense. They could rather be understood as complement-avoiding strategies. Thus, the attributive participles in (99a) and (99b) represent a relativization strategy with the agent participle in *-yś*. The agent of the complement predicate is an argument of the matrix clause and the complement predicate is an attributive participle. In example (100) from Komi, a similar structure is preceded by a *myj*-complement, so one may consider this structure as a strategy to avoid two *myj*-complements following each other ('...quarrel that they forgot that the girl will come today').

(99) a. Komi
Me vunöd-i aski lokt-yś Anna jyv-śy-s
I forget-PST.1SG tomorrow come-PAG Anna about-ELA-3SG
b. Ud.
Mon čukaźe lykt-iś Anna śaryś vunet-i.
I tomorrow come-PAG Anna about forget-PST.1SG
'I forgot about tomorrow coming Anna.' (elicited)

(100) KoZ
Kodyr lokt-isny, kut-isny viťćy-ny, [myj ńija vunöd-isny
when come-PST.3PL begin-PST.3PL quarrel-INF what they forget-PST.3PL
tavun-ja lokt-an nyl-sö].
today-ADV come-PRP girl-ACC.3SG
'When they arrived they started to quarrel that they forgot about the girl which was supposed to come today.'
(Middle Sysola; UK3: 38.3)

Finally, two elliptic strategies have to be mentioned. The first is represented by the more or less regular omission of the CTP 'wish' in Udmurt examples like (101). The second concerns ellipsis of a complement location predicate of a utterance CTP, accompanied by a raising pattern (subject > object) and a switch from locative to separative coding: in (102) from Komi, the location adverbial is not in inessive ('tells that it *is* in the storehouse'), but in elative ('tells it *from* the storehouse').

(101) Ud. (elicited)
[*Mon uzyr luy-sal!*]
I rich be-COND
'(I wish) I was rich!'

(102) KoZ
“Where is your father’s gun?’
D’eťĭna *viśtal-as* [*ambar-yś*].
boy tell-FUT.3SG storehouse-ELA
‘The boy tells it is in the storehouse’ (lit.: tells it from the storehouse).
(FF II: 1219)

3 Complementizer combination

In this section, some of the complementizers which have been introduced in Section 2 are examined in relation to their combinability. In Section 3.1 pleonastic combinations of general complementizers are examined, in Section 3.2 combinations of general complementizers with elements which contribute special semantics, and in Section 3.3 the combinability of non-finite complements with WH words are discussed.

3.1 Pleonastic combinations

As mentioned in Section 2.1.2, in Komi substandard varieties, the Komi fact-type complementizer *myj* and its borrowed equivalent *što* may appear as a complex conjunction *što-myj* ‘that’; recall (24) repeated here as (103).

(103) KoZ
Me *i-g* *dumajt,* [***yštö-myj*** *zerm-as*]
I NEG.PST-1 think.CN **COMP** rain-FUT3SG
‘I did not think that it would rain.’
(Vym; Žilina 1998: 64, Popova 2011b: 72)

Unlike in Komi, in Udmurt the pleonastic combination of the general complementizer *šuysa* and borrowed *što* leads to the “framing usage” (Ru. *ramočnoe upotreblenie*, Kel’makov 2011) which is due to the two different syntactic positions of the complementizers: one sentence-initial and the other sentence-final. Both elements are considered to contribute the same function, and, consequently, Šutov (1999: 45–46) treats this combination as mere pleonasm. This phenomenon is clearly substandard; Winkler (2001: 75, 2011: 165) describes the double use of Russian and Udmurt conjunctions in the same sentence, as in (104a) and (104b), as not unusual in colloquial Udmurt. Tánczos (2013) examines on the basis of these framing usages the syntactic change from sentence final to sentence initial conjunctions in Udmurt, and, more generally, from OV to VO. She considers double

conjunctions a transitional stage in this change (Tánczos 2013: 111). Besides *što… šuysa*, there are also *jesli… ke* 'if… if' – cf. (62b) above for an example –, or *xotja… ke no* 'though… though', which involve the borrowed Russian conjunctions *esli* 'if' and *xotja* 'though'; other examples of framing are *toľko tunne gine* 'only today (lit. only today only)', or *poka jyryd bygyľes na dyrja* 'as long as your head is still round (lit. as long … as long)' where the first words (*toľko* 'only', and *poka* 'as long as') are, again, Russian borrowings (Kel'makov 2011). Kel'makov (2011: 32) observes that frame constructions with words of Russian origin are typical for the colloquial style of Udmurt, but in early clerical texts translated into Udmurt, Persian and Tatar words borrowed from Tatar prevail.

(104) a. Ud.

Mon	*tod-iśko,*	[***što***	*so*	*lykt-oz*	***šuysa***].
I	know-PRS.1SG	**COMP**	(s)he	come-FUT.3SG	**COMP**

'I know that he will come.'
(Winkler 2001: 75, 2011: 165)

b. Ud.

So	*malpa-s,*	[***što***	*kreśan*	*so-ly*	*göršok-yn*	*zarńi*
(s)he	think-PST.3SG	**COMP**	peasant	(s)he-DAT	pot-INE	gold
vaj-em	***šuysa***].					
bring-PRF	**COMP**					

'He thought that the peasant has brought him a pot of gold.'
(OU1: 55.5)

It was demonstrated Section 2.1.1 that *šuysa* is also a purposive conjunction. Example (105) from Beserman contains again a framing construction, but now with *šuysa* in its purposive meaning. Differing from the preceding examples, the sentence-initial slot is occupied here not by a borrowed element, but by the genuine optative particle *med*. Note that the sentence initial position of this particle could not be observed in optative complements; a tendency for this particle to develop into an optative complementizer, as in Komi, was not confirmed (cf. Section 2.1.6 above).

(105) Ud.

Otən	*kək*	*ćas-lə*	*keľt-ono*	*uk,*	***med***	*pe*	*vir-ez*
there	two	hour-DAT	lie-CVB:NEC	MOD	**in.order**	**QUOT**	blood-3SG
dugd-oz		*vaśk-em-əś*	***šusa***.				
stop-FUT.3SG		flow-PP-ELA	**in.order**				

'It's necessary to lie there for two hours in order to stop the blood running.'
(Beserman; Serdobol'skaja et al. 2012: 435)

3.2 Combinations with general complementizers

The Udmurt complementizer conjunction *šuysa* (see Section 2.1.1 above) does not contribute a specific meaning to the sentence and is often optional; Winkler (2011: 166) calls it a facultative indicator of subordination. It always occurs in sentence final position, thus leaving open the sentence-initial slot for another, possibly conjunctional, element, either borrowed or genuine (cf. also Section 3.1). This pattern is attested, for example, with WH-word complements as in (106a) and (106b). The combination of the sentence initial *kyźy* 'how' and the complementizer *šuysa* in immediate perception complements is exemplified by (106c). This combination seems to be rare, probably because *kyźy* as an indicator of state-of-affairs (non-truth-valued) complements is often not compatible with a marker of propositional (truth-valued) complements.

(106) a. Ud.
So *jua-z,* [***kin-o*** *bert-iz* ***šuysa***].
(s)he ask-PST.3SG **who-ADVS** come-PST.3SG **COMP**
'He asked who did actually come?'
(Winkler 2011: 171)

b. Ud.
Soku *tod-em* *ińi,* [***kine*** *so-ly* *śot-ono* ***šuysa***].
then know-PRF.3SG already **who.ACC** (s)he-DAT give-PNEC **COMP**
'At that time he knew already who shall be given to him (as his wife).'
(Winkler 2011: 171, 166)

c. *Anna* *śinjylt-iz,* [***kyźy*** *pijaš* *gožtet-se* *gožt-e* ***šuysa***].
Anna observe-PST.3SG **how** boy letter-ACC3SG write-PRS.3SG **COMP**
'Anna observed how the boy wrote a letter.' (elicited)

The combination of *šuysa* with non-finite forms was not attested in my data. But Serdobol'skaja et al. (2012: 401) report the combination of *šuysa* with the infinitive in purposive adjuncts as a construction occurring in the Beserman dialect of Udmurt.

The 'what'-complementizers, Komi *myj*, and the borrowed *što* in all Permic languages, combine with the conditional word *jeśĺi* (see Section 2.2.3 above), with the quotative particle *pö* (see Section 2.5 above), and with the optative particle *med(ym)*. This last combination is exemplified by (107) with the CTP 'fear'. The complement-internal negation in this example is triggered by the fact that the state of affairs expressed in the complement is undesirable, a typical phenomenon in Komi dialects, according to Serdolbol'skaja et al. (2012: 431) who report it for Pečora and Ižma. It seems that this negation is dependent on the use of the particle *med*; cf. (108a) vs. (108b)–(108c). All three constructional variants are considered synonymous by my Komi consultant, although the complement

negation and the particle *žö* appear only in (108b) and (108c). The order of the complementizer and the optative particles in (108c) is fixed.

(107) KoZ
I *ać-ym* *tyrź-a,* *sečöma* *pol-a,* [***myj*** ***med*** *miśa*
and self-1SG shiver-PRS.1SG so.much fear-PRS.1SG **COMP** **COMP.OPT** AUTQ
köť *ošk-ys* *ješö* ***o-z*** ***veťćy*** *me* *bör-yś*].
yet bear-3SG even.more **not.PRS-3** **follow.CN** I after-ELA
'And I shiver, so much am I afraid that, as I think, the bear may even follow me.'
(Middle Sysola; UK3: 54.4)

(108) KoZ (elicited)
a. *Me* *pol-a,* [***myj*** *ošk-ys* *vötć-as* *me* *börśa*].
I fear-PRS.1SG **COMP** bear-3SG follow-FUT.3SG I after
b. *Me* *pol-a,* [***med*** *ošk-ys* ***o-z*** ***žö*** ***vötćy*** *me* *börśa*].
I fear-PRS.1SG **COMP.OPT** bear-3SG **not.PRS-3** **MOD** **follow.CN** I after
c. *Me* *pol-a,* [***myj*** ***med*** *ošk-ys* ***o-z*** ***žö*** ***vötćy***
I fear-PRS.1SG **COMP** **COMP.OPT** bear-3SG **not.PRS-3** **MOD** **follow.CN**
me *börśa*].
I after
'I'm afraid the bear will follow me.'

Another combining element is the borrowed similative/epistemic particle *byťťö(kö)* (see Section 2.1.7); cf. example (109). The example comes from a parallel text, and interestingly, the element *byťťökö* occurs only in the Komi text, but not in the Russian original, where we find only *čto* 'that'. In our opinion, this shows that the use of *byťťö(kö)* in marking speaker's opinion on reported information, is used in Komi independently from Russian. The use of the general complementizer *myj* with this particle is facultative as (110a) shows, although the order of the two elements is fixed. (110a) and (110b) contain also the quotative particle *pö* which adds the reportative meaning. While the additional use of *myj* in (110a) does not influence the epistemic character of the complemented sentence, the use of *myj* without *byťťökö* in (110b) lacks this meaning.

(109) KoZ
Finnland'ija-yn *i* *18.* *nem-as* *na* *una-ön* *ziľ-isny*
Finland-INE also 18th century-INE.3SG still many-INS engage-PST.3PL
petködly-ny, [***myj*** ***byťťökö*** *suomi* *kyv* *rödvuž-a* *ivrit,*
show-INF **COMP** **as.if** Finnish language related-ADJ Hebrew
važ *grećesköj* *da* *gotsköj* *kyv-jas-köd*].
Ancient Greek and Gothic language-PL-COM
'In Finland in the 18th century, still many engaged in showing that Finnish is related to Hebrew, Old Greek and Gothic.'
(Suomi 493)

(110) KoZ (elicited)

a. *Marko* *kyvl-öma,* [*myj]* ***byťťö*** *jen* *vol-öma* *pö*].
Marko hear-PRF.3SG COMP **as.if** god be.around-PRF.3SG QP
'Marko has heard that God is said to have been around.'

b. *Marko* *kyvl-öma,* [***myj*** *jen* *vol-öma* *pö*].
Marko hear-PRF.3SG **COMP** god be.around-PRF.3SG QP
'Marko has heard that God was around.'

3.3 Combinations of WH words and non-finite complements

It seems that Udmurt does not allow combining participial complements and 'how'-complements for the expression of immediate perception; cf. (111a) and (111b). The structure *kyźy* + participial complement occurs only if *kyźy* functions as a question word proper, as in (112). In Komi, structures which involve a question word and a non-finite complement are, in general, excluded. Serdobol'skaja et al. (2012: 426) consider them to be specific to Udmurt and Mari.

(111) Ud. (elicited)

a. *Anna* *kyl-e,* [***kyźy*** *pijaš* *gožtet-se* *gožt-e*].
Anna hear-PRS.3SG **how** boy letter-ACC.3SG write-PRS.3SG
'Anna hears how the boy is writing the letter.'

b. *Anna* [*pijaš-leś* *gožtet* ***gožja-m-ze***] *kyl-e.*
Anna boy-ABL letter **write-PP-ACC.3SG** hear-PRS.3SG
'Anna hears how the boy is writing a letter.'

(112) Ud. (elicited)

Mon *vunet-i* [*Anna-leś* ***kyźy*** *lykty-ny* ***ödja-m-ze***].
I forget-PST.1SG Anna-ABL **how** come-INF **want-PP-ACC.3SG**
'I have forgotten how [e.g. by which means of transport] Anna wanted to come.'

4 Bifinite vs. monofinite complementation: complexity and information structure

As was stated in the introduction, there are competing constructions for fact-type and immediate perception complements, and in Udmurt, for WH-complements as well: a finite construction with a complementizer ('what', 'how', WH, Ud. *šuysa*) vs. a participle inflected for case and person. This competition of structures is, naturally, not restricted to complementation; it is, in fact, a general phenomenon of subordination in Permic, as well as in many other languages. In a study on competing embedding structures in temporal subordination in Komi,

Leinonen (1998: 96) states that "conjunctions followed by subordinate clauses, as compared to synthetic verbal constructions, are simpler and thus require less planning in actual speech production," as well as that "conjunctions may also serve as signals of continuation indicating to the listener that the message is not finished yet; they let the speaker keep the turn in conversation." These are to be understood as advantages of conjunctional over synthetic syntax, which explains why the conjunctional structures – as demonstrated by Leinonen's investigation – are generally more frequent. Apart from this observation on favouring conditions for the use of conjunctions one must also consider at least two more factors: (i) syntactic complexity and (ii) information structure. First, the capability of the participle predicate in taking arguments and adverbs is different in Komi and Udmurt. While in Udmurt a participial phrase may be rather complex, in Komi it may not be. This is probably the reason why the purposive clause taking fact type complement 'Danila often scratches with a razor under his nose in order for his moustache to grow better' in (113) is expressed with a finite conjunction clause in Komi-Zyrian and Komi-Permyak while the Udmurt variant shoulders a participle.

(113) a. KoZ
Pavel töd-ö, [*myj Dańila ćastö vušt-ö britva-ön nyr*
Pavel know-PRS.3SG COMP Danila often scratch-PRS.3SG razor-INS nose
uv-sö, medym usk-ys bydm-is bur-džyk-a].
under-ACC.3SG in.order moustache-3SG grow-PST.3SG good-CMP-ADV

b. KoP
Pavel töd-ö, [*što Dańilo nyr uvt-sö ćasto*
Pavel know-PRS.3SG COMP Danila nose under-ACC.3SG often
zyrt-ö britva-ön, medby bur-žyk-a bydm-isö us-sez].
scratch-PRS.3SG razor-INS in.order good-CMP-ADV grow-PST.3PL moustache-PL

c. Ud.
[*Myjyk-ez umoj med bud-oz šuysa, Dańila-leś*
moustache-3SG good OPT grow-FUT.3SG in.order Danila-ABL
nyr-ul-yśty-z myjyk-se britva-en čem vuštyl-em-ze]
nose-under-ELA-3SG moustache-ACC.3SG razor-INS often scratch.FRQ-PP-ACC.3SG
Pavel tod-e.
Pavel know-PRS.3SG
'Pavel knows that Danila often scratches with a razor under his nose in order for his moustache to grow better.'
(PM 191)

Secondly, according to our observations, the choice between a finite complement with a clause initial (Ko. *myj*) or clause final (Ud. *šuysa*) complementizer conjunction and a non-finite participial complement is, at least partly, determined by information structure. Note that a participial complement is formally a possessive NP, cross-referencing the agent with a possessive suffix (in Udmurt, not in Komi, see Section 2.3.2 above). This possessive relationship, obviously, prefers

given information, i.e. participial complements tend to convey given information, while new information is expressed via finite complements. This is illustrated by example (114), in which an established situation is continued in three different versions. All three versions contain a predication with the verb 'see' and its complement. Only in the first version does the complement contain old information, repeating a part of the situation established before. In the other two versions, the complement clause contains new information (a new manner adverb, or a completely new situation). For both Permic languages consultants produced automatically (Udmurt) or accepted readily (Komi) a participial complement for the first continuation with old information, but not for the other two with new information – here both insisted that a finite subordinate clause should be used.

(114) Context: When looking out of the window Olga saw how Anna left the house and ran towards the forest. (elicited examples)
Continuation A:
a. KoZ
Oľga *aďďž-is* [*Anna-lyś* *kotört-öm-sö*] *da* *pyr* *žö*
Olga see-PST.3SG Anna-ABL run-PP-ACC.3SG and immediately also
mun-is *bör-śań-ys*.
go-PST.3SG after-EGR-3SG
b. Ud.
[*Anna-leś* *byźy-sa* *pot-em-ze*] *addžy-sa,* *Oľga* *so-len*
Anna-ABL run-CVB go.out-PP-ACC.3SG see-CVB Olga (s)he-GEN
bör-śaz *pot-iz*.
after-EGR.3SG go.out-PST.3SG
'Olga saw Anna running and went after her immediately.'
Continuation B:
a. KoZ
Oľga *aďďž-is,* [*myj* *Anna* *kotört-ö* *zev* *ödjö*], *da* *sijön*
Olga see-PST.3SG COMP Anna run-PRS.3SG very quickly and therefore
e-z *mun* *bör-śań-ys*.
not.PST-3SG go.CN after-EGR-3SG
b. Ud.
Oľga *addž-iz,* [*Anna* *tuž* *džog* *byź-e* *šuysa*], *sojin*
Olga see-PST.3SG Anna very quickly run-PRS.3SG COMP therefore
bör-śaz *ö-z* *poty*.
after-EGR.3SG not.PST-3SG go.out.CN
'Olga saw that Anna was running very quickly and therefore did not go after her.'
Continuation C:
a. KoZ
Oľga *aďďž-is,* [*myj* *lokt-ö* *śöd* *zer-a* *kymör*], *da*
Olga see-PST.3SG COMP come-PRS.3SG black rain-ADJ cloud and
Anna-ys *lo-i* *sy-ly* *žal'*.
Anna-3SG become-PST.3SG (s)he-DAT sorry

b. Ud.
Oľga [*gudyrja-sa* *zor* *lykt-e* *šuysa*] *adďž-iz* *no*
Olga thunder-CVB rain come-PRS.3SG COMP see-PST.3SG and
Anna-jez *žaľa-ny* *kutsk-iz.*
Anna-ACC pity-INF start-PST.3SG
'Olga saw that a black rain cloud/a thunder storm was coming and started to pity Anna.'

The results from elicitation can be corroborated by data from narrative folklore; consider the examples in (115) from a narrative in the Middle-Sysola dialect of Komi. The participial complement in (115a) conveys old information: the approaching of the young man has been mentioned immediately before. Differently, the paratactic finite complement in (115b) conveys new information. And in the Udmurt example (116a), the participle of the CTP 'understand' contains the given information of grandfather's coming, whereas the subordinated clause complement in (116b) consists of new information.

(115) KoZ
a. *Kodyr* *sijö* *nyl-ys* *aďďž-is* [*lokt-öm-sö* *tom* *mort-ys-lyś*],
when this girl-3SG see-PST.3SG come-PP-ACC.3SG young man-3SG-ABL
vištal-is *baťa-ys-ly...*
say-PST.3SG dad-3SG-DAT
'When this girl saw that the young man was coming she told her father ...'
b. *Aďďž-is,* [*lokt-öny* *vör-śań-ys* *jöz*].
see-PST.3SG come-PRS.3PL forest-EGR-3SG people
'She saw people coming from the forest.'
(Middle Sysola; UK3: 38.7, 36.14)

(116) Ud.
a. [*Śerjoga* *peśataj-ez-leś* *maly* *lykt-em-ze*] *pijaš*
Sergej grandfather-3SG-ABL why come-PP-ACC.3SG boy
vala-z *no* *köt-yz* *džož* *lu-iz.*
understand-PST.3SG and stomach-3SG sad become-PST.3SG
'The boy understood why grandfather Sergej had come, and became sad.'
b. *Jakov* *ki-os-yn-yz* *kyče-ke* *vyros-jos* *leśt-iz* *no,* *so-os*
Jakov hand-PL-INS-3SG kind.of-INDF move-PL make-PST.3SG and (s)he-PL
[*bakćaj-e* [...] *myny-ny* *kos-e* *šuysa*] *vala-zy.*
garden-ILL go-INF want-PRS.3SG COMP understand-PST.3PL
'Jakov made some moves with his hands, and they understood that he wants them to go to the garden, behind the shed.'
(PM 903, 1004)

In Udmurt, however, the situation seems to be less straightforward than in Komi. In Udmurt narratives, there are also many examples in which a participial complement contains new information; cf (117) which is the first sentence of a narrative. It is, however, necessary to state that the information structural difference investigated here differs from the idea that only finite complements would

express propositions with a truth-value, while participial complements would express state-of-affairs without a truth-value. The sentence with the participial complement in (117) was continued with the sentences in (118). The Udmurt consultant had no difficulties in accepting these continuations in which the proposition from the complement sentence in (117) is either formally questioned (118a), epistemically evaluated (118b), or denied (118c). This test shows that, at least in present-day Udmurt, non-finite structures such as participles are not reduced predicational options.

(117) Ud.
Mama-jet *šur* *dur-yn* *gyljaśk-e* *vyl-em* *i* [*kion-leś* *džaźeg-ez*
mom-2SG river bank-INE rinse-PRS.3SG be-PRF.3SG and wolf-ABL goose-ACC
nu-em-ze] *aďď-em.*
carry-PP-ACC.3SG see-PRF.3SG
'Your mother was rinsing at the riverside and saw a wolf taking away a goose.'
(OU2: 225.6)

(118) Ud. (elicited)
a. *So* *aďžy-m-te,* [*kytćy* *so* *so-je* *nuy-sa* *košk-em*].
(s)he see-PP-ABE whereto (s)he (s)he-ACC bring-CVB go.away-PRF.3SG
'She did not see where he took it.'
b. *So* *u-g* *tody* *šońer,* [*nu-em* *ik-a* *kion* *džaźeg-ez,*
(s)he not.PRS-3 know.CN really carry-PRF3SG MOD-QP wolf goose-ACC
jake *nuy-m-te*].
or carry-PP-ABE
'She did not see whether the wolf took the goose really away.'
c. *So* *jangyš* *aďď-em.* *Kion* *töďy* *čuškon-ez* *nu-em.*
(s)he wrong see-PRF.3SG wolf white towel-ACC carry-PRF.3SG
'She was mistaken, the wolf carried away a white towel.'

5 Conclusions

This paper presents an overview of complementation strategies in the Permic languages Komi (Zyrian and Permyak), and Udmurt. In many respects, it has more of a preliminary character since many questions could not be answered exhaustively for each of the Permic languages. However, it was shown that while those morphemes that mark non-fact type sentence mood of a finite complement clause are usually the same as in independent clauses (WH-words, question particles, epistemic particles, conditional particles) for imperative and optative complements, there is a special complementizer in Komi, *med(ym)*, which mirrors the Russian *čtoby*, which is in turn is borrowed in many varieties of Permic. In Udmurt, the optative particle *med* has not (yet) developed into a complementizer. Those mor-

phemes which acquire the function of indicating finite complement clauses are thus either of Russian origin – *čto*, *čtoby*, eventually *budto* – , or calques from Russian – Komi *myj*, eventually *med(ym)*. In order for a native morpheme to develop into a complementizer conjunction it must be able to move to the initial position of the subordinate clause. This, however, is impossible for the enclitic elements – question, conditional, and quotative particle. In the case of Udmurt *šuysa* the sentence-final position is the relevant one, but the variety of functions of this word is too broad to call it a fact-type complementizer, like Komi Zyrian *myj*, or Old Zyrian *veveś*; it is also a purposive conjunction – like Russian *čtoby*, and Komi *med(ym)* – and appears with optative complements.

It also turned out that it is difficult to present an exact account of modality in embedded clauses because the syntax of the languages in question is not sufficiently described, and all three languages seem to have several overlapping patterns of which some are older, genuine structures, others clearly Russian, and others adaptations of (more or less) lately transferred Russian patterns. In addition, interaction with modal particles has not been investigated yet.

Besides finite complements with or without complementizer conjunctions, case-inflected participial complements are important; object sentence complements, inflected for the accusative, make the largest group. Again, Permic languages differ in several respects. Komi does not mark the subject on the participle, and the syntactic possibilities of Komi participle are restricted – but not as restricted as of the infinitive. For Udmurt, however, it must be stated that they are functionally in no way restricted, and able to convey also propositional meaning. They are not – as Serdobol'skaja et al. (2012: 462) summarize – necessarily expressing events rather than facts. Furthermore, the variation between finite and participial complements is governed also by at least two more parameters. The first is information structure: if the content of the complemented sentence is given, the perfect participle may be employed, if it is new, the bifinite strategy with a complementizer is preferred. The second is complexity: if the complement predication is (too) complex, a finite strategy is employed.

In conclusion one may say that the possible semantic impact of those morpho-syntactic means that identify a proposition as a complement in Permic languages is quite small, but growing.

Acknowledgment

Foremost thanks go to Anna Baidullina and Nikolai Kuznetsov in Tartu for their expertise and insistence; as well as to Kasper Boye and Petar Kehayov for initiat-

ing the present volume, as well as substantial discussion and comments. Further thanks go to Jorma Luutonen at the Research Unit for Volgaic Languages at University of Turku for enabling access to data, to Inna Kaysina (Bremen) and to an anonymous reviewer for valuable comments. This research was supported by the *European Social Fund's Doctoral Studies and Internationalisation Programme DoRa*.

References and Sources

Bartens, Raija. 2000. *Permiläisten kielten rakenne ja kehitys* [Structure and development of Permic languages] (Mémoires de la Société Finno-Ougrienne 238). Helsinki: Suomalais-Ugrilainen Seura.

Belyh, Sergej K. 2002. The disintegration of the Proto-Permian community: On the history of studies. *Finnisch-Ugrische Mitteilungen* 24/25. 29–53.

Biryuk, Olga & Maria Usačeva. 2011. Russian in Beserman oral discourse: Global influence and interaction. In Sándor Csúcs, Nóra Falk, Viktória Püspök, Viktória Tóth & Gábor Zaicz (eds.), *Congressus XI Internationalis Fenno-Ugristarum, Piliscsaba 2010,* Pars V, 258–266. Piliscsaba: Reguly Társaság.

Brykina, Marija M. & Natalija. B. Aralova. 2012. Sistemy pričastij v marijskom i permskix jazykax [The participle systems of Mari and Permic]. In Ariadna I. Kuznecova (ed.), *Finno-ugorskie jazyki: fragmenty grammatičeskogo opisanija. Formal'nyj i funkcional'nyj podxody* [Finno-Ugric languages: fragments of a grammatical description. Formal and functional approaches] (Studia Philologica), 476–520. Moscow: Rukopisnye pamjatniki drevnej Rusi.

Bubrix, Dmitrii V. 1949. *Grammatika literaturnogo komi jazyka* [Grammar of literary Komi]. Leningrad: Nauka.

Csúcs, Sándor. 1998. Udmurt. In Daniel Abondolo (ed.), *Uralic Languages*, 276–304. London & New York: Routledge.

Cypanov, Evgenij A. 1997. *Pričastie v komi jazyke: istorija, semantika, distribucija* [Participle in Komi: history, semantics, distribution]. Ekaterinburg: UrO RAN.

Cypanov, Evgenii A. 2005. *Grammatičeskie kategorii glagola v komi jazyke* [Grammatical categories of the Komi verb]. Syktyvkar: Komi Naučnyj Centr UrO RAN.

Cypanov, Evgenij A. 2010. Ilja Vas' da komi neologijaly poduvyas [Ilja Vas' and the foundation of Komi neology]. In Elena V. Ostapova (ed.), *V. I. Lytkin: grani nasledija* [V. I. Lytkin: limits of legacy], 29–34. Syktyvkar: Syktyvkar State University.

FF I–II = Fokos-Fuchs, David R. 1959. *Syrjänisches Wörterbuch*. I–II. Budapest: Akadémiai kiadó.

Geisler, Michael. 2005. *Vokal-Null-Alternation, Synkope und Akzent in den permischen Sprachen* (Veröffentlichungen der Societas Uralo-Altaica 68). Wiesbaden: Harrassowitz.

GSUJa. 1962 = Perevoščikov, Petr N. (ed.). 1962. *Grammatika sovremennogo udmurtskogo jazyka. Fonetika i morfologija* [Grammar of contemporary Udmurt: phonetics and morphology]. Iževsk: Udmurtskoe knižnoe izdatel'stvo.

GSUJa. 1970 = Vaxrušev, Vasilij M., Vasilij N. Zaxarov & L. I. Kalinina (eds.). 1974. *Grammatika sovremennogo udmurtskogo jazyka. Sintaksis složnogo predloženija* [Grammar of contemporary Udmurt. Syntax of the complex sentence]. Iževsk: Udmurtija.

Hamari, Arja. 2011. The abessive in the Permic language. *Journal de la Société Finno-Ougrienne* 93. 37–84.

Hansen, Björn, Alexander Letuchiy & Izabela Błaszczyk, this volume. Complementizers in Slavonic (Russian, Polish, and Bulgarian).

Hausenberg, Anu-Reet. 1998. Komi. In Daniel Abondolo (ed.), *Uralic Languages*, 305–326. London & New York: Routledge.

Honti, László. 1994. Slawischer Einfluss auf die finnisch-ugrischen Sprachen. *Incontri Linguistici* 17. 81–101.

Jašina, Roza I. 1981. Sintaksičeskie konstrukcii so slovom *šuysa* v udmurtskom jazyke [Syntactic constructions with the word *šuysa* in Udmurt]. In *Istoriko-kul'turnye svjazi permskix narodov: Po dannym fol'klora i jazyka* [Historical-cultural relations of Permic peoples: according to data from folklore and language], 78–95. Iževsk: Udmurt State University.

Karpova, Ljudmila L. 1997. *Fonetika i morfologija srednečepeckogo dialekta udmurtskogo jazyka* [Phonetics and morphology of the Middle Cheptsa dialect of Udmurt] (Dissertationes Philologicae Uralicae Universitatis Tartuensis). Tartu: Tartu Ülikooli Kirjastus.

Kel'makov, Valentin K. 1973. Ob otglagol'nyx poslelogax *šuysa* i *luysa* v udmurtskom jazyke [On the deverbal postpositions *šuysa* and *luysa* in Udmurt]. In Vasilij M. Vaxrušev (ed.), *Voprosy udmurtskogo jazykoznanija* [Questions of Udmurt linguistics] II, 123–151. Iževsk: Udmurtskij naučno-issledovatel'skij institut istorii, ėkonomiki, literatury i jazyka.

Kel'makov, Valentin. 2011. Ramočnoe upotreblenie sinonmičnyx služebnyx slov v udmurtskom jazyke (v kontekste jazykov Uralo-Povolzh'ya) [Framing use of synonymous functional words in Udmurt [in the context of the Ural-Volga area]]. In Sándor Csúcs, Nóra Falk, Viktória Püspök, Viktória Tóth & Gábor Zaicz (eds.), *Congressus XI Internationalis Fenno-Ugristarum, Piliscsaba 2010*, Pars IV, 215–224. Piliscsaba: Reguly Társaság.

KĖSK = Lytkin, Vasilij I. & Evgenij S. Guljaev. 1999. *Kratkij ėtimologičeskij slovar' komi jazyka* [Short etymological dictionary of Komi]. Syktyvkar: Komi knižnoe izdatel'stvo.

Klumpp, Gerson. 2009. Variation in Komi object marking. In Andreas Dufter, Jürg Fleischer & Guido Seiler (eds.), *Describing and modeling variation in grammar* (Trends in Linguistics, Studies and Monographs 204), 325–359. Berlin & New York: Mouton de Gruyter.

Klumpp, Gerson. 2014. Identifiability, givenness and zero-marked referential objects in Komi. *Linguistics* 52(2). 415–444.

KRS = Beznosikova, Ljucija M., Evgenija A. Ajbabina & Raisa I. Kosnyreva. 2000. *Komi-russkij slovar' – Komi-roč kyvčukör* [Komi-Russian dictionary]. Syktyvkar: Komi knižnoe izdatel'stvo.

L52 = Lytkin, Vasilij I. 1952. *Drevnepermskij jazyk. Čtenie tekstov, grammatika, slovar'* [Old Permic. Text readings, grammar, glossary]. Moscow: Izd. Akademii Nauk SSSR.

L61 = Lytkin, Vasilij I. 1961. *Komi-jaz'vinskij dialekt* [The Jazva dialect of Komi]. Moscow: Izdatel'stvo Akademii Nauk SSSR.

Leinonen, Marja. 1998. Russian influence on sentential subordination in Komi Zyrian. *Dialectologica et Geolinguistica* 6. 81–98.

Leinonen, Marja. 2000. Evidentiality in Komi-Zyrian. In Lars Johanson & Bo Utas (eds.), *Evidentials. Turkic, Iranian and neighbouring languages*, 419–440. Berlin & New York: Mouton de Gruyter.

Leinonen, Marja. 2002. Influence of Russian on the syntax of Komi. *Finnisch-Ugrische Forschungen* 57. 195–358.

Maitinskaja, Klara. 1982. *Služebnye slova v finno-ugorskix jazykax* [Functional words in Finno-Ugric languages]. Moscow: Nauka.

Manova, Nadežda D. 1975. Složnopodčinennye predloženija v komi jazyke s sojuzami nemestoimennogo proisxoždenija [Subordinate clauses in Komi and conjunctions of non-pronominal origin]. *Sovetskoe Finnougrovedenie* 1975(2). 90–95.

Manova, Nagyezsda. 1998. *Komi-zürjén nyelvkönyv* [Komi-Zyrian textbook] (Studia Uralo-Altaica Supplementum 7). Szeged: JATE Finnugor tanszék.

MF52 = Munkácsi, Bernát & David-Rafael Fuchs. 1952. *Volksbräuche und Volksdichtung der Wotjaken* (Mémoires de la Société Finno-Ougriennne 102). Helsinki: Suomalais-Ugrilainen Seura.

MSz. = Munkácsi, Bernát 1896. *A votják nyelv szótára* [Votjak dictionary]. Budapest: Magyar Tudományos Akadémia.

Osnovy 1974 = Lytkin, Vasilij I., Klara E. Majtinskaja & Karoly Rédei (eds.). 1974. *Osnovy finno-ugorskogo jazykoznanija* [Foundations of Finno-Ugric linguistics], vol. 1: *Voprosy proisxoždenija i razvitija finno-ugorskix jazykov* [Questions of origin and development of Finno-Ugric languages]. Moscow: Nauka.

OU1 = Kel'makov, Valentin I. 1981. *Obrazcy udmurtskoj reči. Severnoe narečie i sredinnye govory* [Examples of Udmurt speech. Northern dialect and central varieties]. Iževsk: Naučno-issledovatel'skij institut pri sovete ministrov Udmurtskoj ASSR.

OU2 = Kel'makov, Valentin I. 1990. *Obrazcy udmurtskoj reči II: Sredinnye govory* [Examples of Udmurt speech. Central varieties]. Iževsk: Udmurtija.

Perevoščikov, Petr N. 1959. *Deepričastija i deepričastnye konstrukcii v udmurtskom jazyke* [Adverbial participles and adverbial participle constructions in Udmurt]. Iževsk: Udmurtskoe knižnoe izdatel'stvo.

PM = Pavel Morozov. Parallel text. Corpora of the research unit for Volgaic languages. University of Turku.

Popova, Ėleonora N. 2011a. Služebnaja edinica *med* v permskix jazykax [The functional unit *med* in Permic languages]. In Evgenij A. Cypanov (ed.), *Dinamika struktur finno-ugorskix jazykov* [Dynamics in the structures of the Finno-Ugric languages], 197–205. Syktyvkar: Kola.

Popova, Ėleonora N. 2011b. Sojuzy i časticy v severo-zapadnyx dialektax komi jazyka [Conjunctions and particles in North-Western dialects of Komi]. In Evgenij A. Cypanov, (ed.), *Dinamičeskie processy v sistemax permskix jazykov* [Dynamic processes in the systems of Permic languages] (Trudy IJaLI 68), 71–77. Syktyvkar: Izdatel'stvo Komi Naučnogo Centra UrO RAN.

R78 = Rédei, Károly. 1978. *Zyrian folklore texts*. Budapest: Akadémiai kiadó.

R96 = Rédei, Károly. 1996. *Zyrian texts* (Studia Uralica 8). Wiesbaden: Harrassowitz.

Rédei, Károly. 1978. *Chrestomathia Syrjaenica*. Budapest: Tankönyvkiadó.

Rédei, Károly. 1988. Geschichte der permischen Sprachen. In Denis Sinor (ed.), *The Uralic languages. Description, history and foreign influences* (Handbuch der Orientalistik 8/1), 351–394. Leiden & New York: Brill.

Riese, Timothy. 1998. Permian. In Daniel Abondolo (ed.), *Uralic Languages*, 249–275. London & New York: Routledge.

Saarinen, Sirkka. 2003. Sprachgesetze und Neologismenbildung bei den wolgaischen und permischen Völkern. In Gerson Klumpp & Michael Knüppel (eds.), *Die ural-altaischen Völker. Identität im Wandel zwischen Tradition und Moderne. Vorträge des Symposions der Societas Uralo-Altaica vom 13.–15. Oktober 2002* (Veröffentlichungen der Societas Uralo-Altaica 63), 91–97. Wiesbaden: Harrassowitz.

Serdobolskaja, Natalia M., Anfisa A. Il'evskaya, Sergey A. Minor, Polina S. Miteva, Aleksandra V. Fajnvejc & Natalia S. Matveeva. 2012. Konstrukcii s sentencial'nymi aktantami v finno-ugorskix jazykax [Constructions with sentential arguments in Finno-Ugric languages]. In Ariadna I. Kuznecova (ed.), *Finno-ugorskie jazyki: fragmenty grammatičeskogo opisanija. Formal'nyj i funkcional'nyj podxody* [Finno-Ugric languages: fragments of a grammatical description. Formal and functional approaches] (Studia Philologica), 382–475. Moscow: Rukopisnye pamjatniki drevnej Rusi.

Serdobolskaja, Natalia & Svetlana Toldova. 2011. Grammaticalization of the verb of speech in Finno-Ugric languages. In Sándor Csúcs, Nóra Falk, Viktória Püspök, Viktória Tóth & Gábor Zaicz (eds.), *Congressus XI Internationalis Fenno-Ugristarum, Piliscsaba 2010*, Pars V, 285–292. Piliscsaba: Reguly Társaság.

Serebrennikov Boris A. 1967. O nekotoryx xarakternyx osobennostjax drevnego sintaksisa jazyka komi [On some characteristic pecularities of Old Komi syntax]. *Sovetskoe Finno-Ugrovedenie* 1967(2). 101–106.

Siegl, Florian. 2004. *The 2nd past in the Permic languages – form, function and a comparative analysis from a typological perspective*. Tartu: University of Tartu MA Thesis. http://dspace.utlib.ee/dspace/bitstream/handle/10062/1265/Siegl.pdf?sequence=5&isAllowed=y (accessed 7 October 2015).

SKJa. II = Sel'kov, Nikolaj N. (ed.). 1967. *Sovremennyj komi jazyk II: Sintaksis* [Contemporary Komi II: syntax]. Syktyvkar: Komi knižnoe izdatel'stvo.

Sorvačëva, Valentina A. & Ljucia M. Beznosikova. 1990. *Udorskij dialekt komi jazyka* [The Udora dialect of Komi]. Moscow: Nauka.

Stipa, Günther. 1960. *Funktionen der Nominalformen des Verbs in den permischen Sprachen* (Mémoires de la Société Finno-Ougrienne 121). Helsinki: Suomalais-Ugrilainen Seura.

Suomi. Parallel text. Corpora of the research unit for Volgaic languages. University of Turku.

Šutov, Aleksandr F. 1999. *Gipotaksis v udmurtskom jazyke. Posobie dlja studentov po sintaksisu* [Hypotaxis in Udmurt. Study aid on syntax]. Iževsk: Izdatel'stvo Udmurtskij universitet.

Tánczos, Orsolya. 2013. *Hogy … hogy?* Kettős kötőszók az udmurt mondatban [That… that? Double conjunctions in the Udmurt sentence]. In Klára Agyagási, Attila Hegedűs & Katalin É. Kiss (eds.), *Nyelvelmélet és kontaktológia* [Language theory and contactology] *2*, 95–112. Piliscsaba: PPKE BTK Elméleti Nyelvészeti Tanszék & Magyar Nyelvészeti Tanszék.

Tauli, Valter. 1966. *Structural tendencies in Uralic languages* (Indiana University Publications, Uralic and Altaic Series 17). The Hague: Mouton.

UK1 = Uotila, Toivo E. 1985. *Syrjänische Texte I. Komi-Permjakisch.* Übersetzt und herausgegeben von Paula Kokkonen (Mémoires de la Société Finno-Ougrienne 186). Helsinki: Suomalais-Ugrilainen Seura.

UK2 = Uotila, Toivo E. 1986. *Syrjänische Texte II. Komi-Syrjänisch: Ižma-, Pečora- und Vym-Dialekte.* Übersetzt und herausgegeben von Paula Kokkonen (Mémoires de la Société Finno-Ougrienne 193). Helsinki: Suomalais-Ugrilainen Seura.

UK3 = Uotila, Toivo E. 1989. *Syrjänische Texte III. Komi-Syrjänisch: Luza-Letka-, Ober-Sysola-, Mittel-Sysola, Prisyktyvkar-, Unter-Vyčegda- und Udora-Dialekte.* Übersetzt und herausgegeben von Paula Kokkonen (Mémoires de la Société Finno-Ougrienne 202). Helsinki: Suomalais-Ugrilainen Seura.

UK4 = Uotila, Toivo E. 1995. *Syrjänische Texte IV. Komi-Syrjänisch: Ober-Vyčegda-Dialekt.* Übersetzt und herausgegeben von Paula Kokkonen (Mémoires de la Société Finno-Ougrienne 252). Helsinki: Suomalais-Ugrilainen Seura.

Vilkuna, Maria. 1998. Word order in European Uralic. In Anna Siewerska (ed.), *Constituent order in the languages of Europe* (Empirical Approaches to Language Typology 20), 173–233. Berlin & New York: Mouton de Gruyter.

W01 = Wichmann, Yrjö. 1901. Wotjakische Sprachproben II. *Journal de la Société Finno-Ougrienne* 19(1).

Winkler, Eberhard. 2001. *Udmurt* (Languages of the World/Materials 212). München: Lincom Europa.

Winkler, Eberhard. 2011. *Udmurtische Grammatik* (Veröffentlichungen der Societas Uralo-Altaica 81). Wiesbaden: Harrassowitz.

WU = Uotila, Toivo E. & Yrjö Wichmann. 1942. *Syrjänischer Wortschatz nebst Hauptzügen der Formenlehre* (Lexica Societatis Fenno-Ugricae VII). Helsinki: Suomalais-Ugrilainen Seura.

Žilina, Tat'jana I. 1985. *Luzsko-letskij dialekt komi jazyka* [The Luza-Letka dialect of Komi]. Moscow: Nauka.

Žilina, Tat'jana I. 1998. *Vymskij dialekt komi jazyka* [The Vym dialect of Komi]. Syktyvkar: Prolog.

Erika Körtvély

Modal functions of the complementizer *hogy* 'that' in Hungarian

1 Introduction

The present paper is concerned with the description and semantic analysis of complementation strategies in Hungarian, in particular with complementation constructions containing the complementizer *hogy* 'that'. After presenting a general overview of clause complementation in Hungarian, I will focus on the semantic functions of *hogy*, which is the most frequently used complementizer in this language. As *hogy* is traditionally assumed to be a neutral subordinating conjunction (as opposed to the other complementizers, which have lexical meanings), the investigation seeks to answer the question of whether it can be regarded as a genuine structural element or can (at least to some extent) be associated with modality and, respectively, considered to be a "lexically separate modality marker that happen[s] to occur in a specific syntactic position" (Frajzyngier 1995: 475).

By the term complementizer I will refer to independent, particle-like linguistic elements that function as subordinators and identify a clause as a complement (cf. Noonan 2007: 55). Modality will be understood in a broad sense – as a linguistic realization of the speaker's attitude toward a given fact (including agent-oriented and speaker-oriented modalities, epistemic modality, and the peripheral fields of evidentiality and illocutionary force [cf. Frajzyngier 1995: 476]).

1.1 The Hungarian language

Hungarian is a Uralic language spoken by approximately 13 million people in Central Europe, in and around Hungary. This number of native speakers makes Hungarian the largest Uralic language. Within the language family, Hungarian belongs to the Ugric branch of the Finno-Ugric languages. The closest relatives of Hungarian are Khanty and Mansi, both highly endangered languages spoken in western Siberia. During and after Hungarians' migration to Europe, Hungarian has been subjected to considerable influences of several Turkic languages, followed by Slavic and German influences, affecting both the grammar and the lexicon of the language. Thus, although the basic grammar of Hungarian has pre-

served many archaic Uralic features, the geographical distance and the peculiar history of linguistic contacts makes Hungarian a non-typical Uralic language.

From a typological point of view Hungarian is a relatively transparent agglutinative language with very rich morphology. Nominal cases, number and possessive marking, as well as verbal inflection (in two numbers and three tenses) are expressed predominantly by suffixation; fusion and suppletion are very limited. Hungarian, similarly to other Uralic languages, does not have grammatical gender but, unlike most Uralic languages, has developed an article system. The syntax of Hungarian is characterized by its discourse-configurational nature. The basic word order is SOV or SVO, and although phrases have a rigid internal structure, it is the information structure that determines their position in the sentence.

Hungarian can be considered a highly uniform language with respect to its dialects and social registers. This paper is based on material representing standard Hungarian and example sentences were tested by speakers of a version called "standard colloquial Hungarian".[1]

1.2 Clausal complementation in Hungarian

Hungarian prefers tensed clause embedding, i.e. an analytic construction for clause complementation where an overt or covert complementizer introduces a complement clause containing a finite verb.

(1) a. *Józsi tudja,* [***hogy*** *Anna moziba ment*].
Józsi know:O.3SG **COMP** Anna cinema:ILL go:PST[3SG]
'Józsi knows that Anna had gone to the cinema.'
b. *Józsi hallja,* [***hogy*** *Anna szépen énekel*].
Józsi hear:O.3SG **COMP** Anna nicely sing[3SG]
'Józsi hears that Anna is singing nicely.'

Complement clauses occur most typically in sentence-final position but given they can be moved to the topic position a sentence-initial and a sentence-internal position is also possible.

c. [***Hogy*** *Anna moziba ment*], *Józsi tudja.*
COMP Anna cinema:ILL go:PST[3SG] Józsi know:O.3SG
'Józsi knows about that, that Anna went to the cinema.'
d. *Józsi,* [***hogy*** *Anna moziba ment*], *tudja.*
Józsi **COMP** Anna cinema:ILL go:PST[3SG] know:O.3SG
'Józsi knows about that, that Anna went to the cinema.'

1 I also have used standard contemporary examples from the Hungarian National Corpus (HNC).

The position of the complementizer itself is strictly bound: it has to be placed clause initially, i.e. the complementizer occupies the leftmost position in the complement clause. As described in the literature (Kenesei 1992: 37; Kenesei & de Urbina 1994: 9; den Dikken & Giannakidou 2002: 46), in Hungarian (unlike e.g. in English) even wh-relatives have to be placed to the right of the complementizer; cf. (2a) and (2b).

(2) a. *Tudom,* [***hogy*** ***miért*** *mentél* *el* *moziba*].
know:O.1SG **COMP** **why** go:PST:2SG PREF cinema:ILL
'I know why you have gone to the cinema.'
b. **Tudom,* ***miért*** *mentél* *el* *moziba* ***hogy***.
know:O.1SG **why** go:PST:2SG PREF cinema:ILL **COMP**

While finite clause complementation is usually independent of the semantics of the matrix verb, non-finite complements have more restricted distribution. However, if the subjects of the matrix clause and the complement clause are identical, and, additionally, the matrix predicate has modal meaning, non-finite complementation with an infinitive can be considered to be more natural than finite clause complementation. Examples (3a) and (3b) represent the two basic types of Hungarian non-finite complement clauses: (3a) the unmarked and (3b) the Dative + Possessive marked infinitive construction (where the subject of the matrix clause is in Dative, and Possessive suffixes express its person and number on the nominalized verb form).

(3) a. *Anna* ***alud-ni*** *szeret-ne*.
Anna **sleep-INF** like-COND[3SG]
'Anna would like to sleep.'
b. *Anná-nak* *szabad* *volt* ***alud-ni-a***.
Anna-DAT allowed be:PST[3SG] **sleep-INF-POSS.3SG**
'Anna was allowed to sleep.'

In some cases, non-finite complementation with different subjects in the clauses is also possible, although the number of the matrix verbs which allow this is very limited. The infinitive complementation is only felicitous with perception verbs, such as *lát* 'see' and *hall* 'hear', cf. (4a)–(4c), and nominalizations with action nouns are usually not interchangeable with finite clause complementation; the examples (5a) and (5b) are intelligible but estimated as not felicitous (not acceptable) by the speakers.

(4) a. *Józsi* *látta* *Annát* *moziba* ***menni***.
Józsi see:PST:O.3SG Anna:ACC cinema:ILL **go:INF**
'Józsi saw Anna going to cinema.'

b. *Józsi* *hallotta* *Annát* *szépen* ***énekelni***.
Józsi hear:PST:O.3SG Anna:ACC nicely **sing:INF**
'Józsi heard Anna singing nicely.'
c. **Józsi* *elhitte* *Annát* *szépen* *énekelni*.
Józsi PREF:believe:PST:O.3SG Anna:ACC nicely sing:INF
Intended: 'Józsi believed that Anna sang nicely.'

(5) a. ???*Józsi* *elhitte* *Anna* *moziba* ***menését***.
Józsi PREF:believe:PST:O.3SG Anna cinema:ILL **go:NOMACT:POSS.3SG:ACC**
Intended: 'Józsi believed that Anna went to the cinema.'
b. ???*Józsi* *látta* *Anna* *moziba* ***menését***.
Józsi see:PST:O.3SG Anna cinema:ILL **go:NOMACT:POSS.3SG:ACC**
'Józsi saw that Anna had gone to the cinema.'

1.3 Complementizers in Hungarian

As mentioned above, the most general strategy of clause complementation in Hungarian is tensed clause embedding (Kenesei 1994b: 2) in which a complementizer identifies an embedded finite clause as a complement. In Hungarian there are numerous linguistic elements fulfilling the syntactic requirements for being classified as complementizers (cf. Kenesei 1992, 1994a, 1994b, Kenesei & de Urbina 1994, Bácskai-Atkári 2011, 2013). These elements, just like many other subordination or coordination conjunctions, are characteristically polyfunctional, and as such they are not restricted to occur exclusively as complementizers (see Section 4 below).

Accepting the categorization of Kenesei (1992: 40–43) we can distinguish between simple and complex complementizers. The Hungarian simple complementizers are *hogy* 'that', *ha* 'if', *bár* 'although' and *mint* 'as, than'. *Hogy* 'that' can be considered to be the most general and by far the most frequently used complementizer, whose semantic meaning is highly abstract. *Ha* 'if' has a conditional meaning and is found in conditional contexts. The concessive and the equative-comparative lexical semantics of *bár* 'although' and *mint* 'as, than', respectively, also determine in which context these markers occur; cf.:

(6) a. *Józsi* *szereti,* ***ha*** *Anna* *olvas.*
Józsi like:O.3SG **COMP** Anna read[3SG]
'Józsi likes (it) if Anna is reading.'
b. ***Bár*** *hideg* *van,* *szép* *az* *idő.*
although cold be[3SG] nice the weather
'Although it is cold, the weather is nice.'
c. *Józsi* *jobban* *hegedül,* ***mint*** (*ahogy*) *Anna.*
Józsi better play.violin[3SG] **COMP** so as Anna
'Józsi plays the violin better than Anna.'

Kenesei (1992: 40; see also Kenesei & de Urbina 1994: 9) regards combinations of two simple complementizers, or mergers of a complementizer and other elements (most usually a wh-phrase) as complex complementizers. These elements function as complex conjunctions and can introduce various types of subordinate and coordinate clauses.

One group of such complex markers includes *hogyha* 'if' (*hogy* 'that' + *ha* 'if'), *bárha* 'if only' (*bár* 'although' + *ha* 'if'), *habár* 'although' (*ha* 'if' + *ha* 'although') *minthogy* 'as' (*mint* 'than' + *ha* 'if'), etc. In fact, almost every combination of the simple complementizers yields a formally valid complex complementizer; cf. example (7a). However, the meaning of a given combination restricts its semantic distribution:

(7) a. *hogyha* 'if' (*hogy* 'that' + *ha* 'if')

Józsi	*hallja,*	***hogyha***	*Anna*	*zongorázik.*
Józsi	hear:O.3SG	**if**	Anna	play.piano:REFL/MID.3SG

'Józsi hears if Anna is playing piano.'

Another group are those complex complementizers, which consist of a simple complementizer and one or more additional linguistic elements; the lexical meaning of this element restricts the use of the whole.

b. *mivelhogy* 'since' (*mivel* 'since' + *hogy* 'that')

Mivelhogy	*Józsi*	*fázott,*	*felvett*	*egy*	*kabátot.*
since	Józsi	be.cold:PST[3SG]	put.on:PST[3SG]	a	coat:ACC

'Since Józsi was cold, he put on a coat.'

c. *mintsemhogy* 'rather than' (*mint* 'as, than' + *is* 'also' + *nem* 'not' + *hogy* 'that')

Inkább	*nem*	*eszem,*	***mintsemhogy***	*hazamenjek.*
rather	not	eat:REFL/MID.1SG	**rather.than**	go.home:SBJV:1SG

'I better don't eat but I will not go home.'

Concerning the position of the elements within complex complementizers, it is important to note that in the compounds of the second group the wh-phrase (or the additional particle) tends to precede the "real complementizer" (i.e. the element occurring alone as a complementizer): forms like *mert+hogy* 'since' or *még+ha* 'even if' are well formed in contrast to **hogy+mert* and **ha+még*. In the first group, in which two simple complementizers form a complex one, the position of the component show variation: *bár+ha* and *ha+bár*, in the (slightly conditional) concessive meaning '~if only, ~although', are both in use in contemporary Hungarian, and there is historical evidence that beside the form *hogy+ha* 'if' the complementizer *ha+hogy* has been also widespread until the second part of the 19th century. Concerning the combinations which do not allow variation in the order of their components, the rule of a thumb is that the more neutral (i.e. the more grammaticalized) component follows the less neutral one. Thus, *mint+ha*

and *bár+hogy* are existing combinations, while **ha+mint* and **hogy+bár* are not acceptable. (For details on the order of the constituents of complex complementizers in historical perspective see Bácskai-Atkári 2011, 2012.)

As shown above, Hungarian has numerous particle-like elements which can occur as complementizers, i.e. they fulfill the syntactic requirements of a subordinator identifying a clause as a complement. Although there are no clitic complementizers in Hungarian in Noonan's sense (2007: 55), a clitic element occurring in combination with *hogy* 'that' has to be introduced as a special inflectional marker necessary for the formation of embedded yes/no-questions. The clitic *-e* occurs attached to the finite verb of the yes/no-question and is used most frequently in embedded *hogy*-complement clauses. Concerning its linguistic status Kenesei and de Urbina (1994: 2) point out that the clitic *-e* is of mixed nature combining complementizer and inflectional functions and may be related to the functional category INFL that hosts Q-features. However, as the interrogative clitic is always coexistent with *hogy* in complementation context (the two together make up a complementizer construction), it is reasonable to classify it as an inflectional or semi-complementizer.

(8) *Józsi* *nem* *tudja,* *hogy* *Anna* *elment=e* *moziba*
Józsi not know:O.3SG COMP Anna PREF:go:PST[3SG]=Q(yes/no) cinema:ILL
'Józsi does not know whether Anna has gone to the cinema.'

I will focus in the following sections on the most frequent simple complementizer *hogy* 'that' and its relevant combinations. After a description of the possible semantic functions of *hogy* (Section 2) I describe its distribution and the specific semantic or pragmatic features associated with it (Section 3). In Section 4 I deal with the issue of complementizer omission and in Section 5 I provide an overview of the non-complementizing functions of this neutral complementizer which show that despite sharing the general characteristics of the word class of complementizers it is a linguistic element with more than a single function. Section 6 provides information about the diachronic development and grammaticalization of *hogy*. In Section 7 the most important findings of the paper are summarized.

2 The complementizer *hogy* 'that'

2.1 That-complementizers and if-complementizers?

Hungarian does not have the straightforward distinction between 'that-complementizers' and 'if-complementizers' we encounter in other languages (cf. e.g. Huddleston & Pullum 2002: 956 and Nordström 2010: 9, 91). Although there is a complementizer with the meaning 'if' in Hungarian, its distribution does not correlate with the complementizers 'if' and 'weather' of the Germanic languages for example. The Hungarian *ha* 'if' has a clear conditional meaning and as such does not show contrast to the neutral *hogy* 'that' in the way the Germanic conditional/interrogative complementizers are contrasted with the neutral (declarative) complementizers. Furthermore, as was already shown in Section 1.3, *hogy* and *ha* usually do not exclude each other but (combined with each other or with other elements) form complex complementizers.

Nevertheless, the semantic opposition between 'that-complementizers' and 'if-complementizers' is present also in Hungarian, although it can be associated with certain syntactic structures in *hogy*-complementation rather than with a distinct complementizer form.

2.2 COMPs, complementizers, and subordinators

Complementation and complementizers have been an object of increasing attention in the last two decades, but due to the different research questions, methods and theoretical platforms, there is still little agreement in the field what these terms actually designate. Therefore, before analyzing the possible semantic functions of *hogy*, we have to clarify its linguistic status.

The existing research on complementation and complementizers in Hungarian (Kenesei 1992, 1994b; Kenesei & de Urbina 1994; Szabolcsi 1992; Kiss 2002; Kenesei, Vágó & Fenyvesi 1998) regards *hogy* and other complemetizers (COMPs) as structural elements the function of which is to fill the complementizer position in a clause. As these works are predominantly structurally oriented, complementizers are treated in them as special subordinators, irrespective of their actual meanings or semantic functions.

However, as observed by many authors (e.g. Bolinger 1968, Kirsner & Thompson 1976, Frajzyngier & Jaspersen 1991: 133; Frajzyngier 1995: 476), complementizers often contribute to the expression of modality in the language and as such are more than merely syntactic elements. Bhatt and Yoon (1992) offer the following

terminological solution to this double nature of complementizers. They consider complementizers as a non-unified category in natural languages, both functionally and structurally (Bhatt & Yoon 1992: 41). Due to the two distinct functions (identification of the clause type and modal functions) such elements have, a distinction between (lexical) complementizers and subordinating elements has to be made. Thus, subordinators carry only information concerning the subordinate status of a structural entity, whereas lexical complementizers serve as subordinators on the one hand and indicate clause types and mood on the other. In a large number of languages (e.g. in Germanic languages) lexical complementizers convey both functions simultaneously but there are also languages in which these functions are formally kept apart. As Bhatt and Yoon point out, usually "languages with a robust agglutinative morphology" (Bhatt & Yoon 1992: 42) have separate means of expressing modality in complement clauses. At the same time, the comparison between complementizer systems of languages with subordinator complementizers and complementizers systems of languages with lexical complementizers is problematic, because a neutral complementizer like the English *that* (opposed to *if* and *whether* in its own complementizer system) cannot be treated as equivalent to a neutral complementizer like the Hungarian *hogy* 'that' (opposed to no complementizer with the semantics of 'if' and 'whether' in its own system).

However, as I aim to show below, the complementizer *hogy* is not as modality-insensitive as it would follow from the fact that Hungarian does not have complementizers with the meaning and function of English *if* and *whether*.

2.3 Complement clauses and modality in Hungarian

The exclusively syntactic focus of the research on complementation in Hungarian could be blamed for the fact that there are virtually no studies on the possible semantic functions of the complementizers. Furthermore, even semantic research (e.g. Tóth 2006, 2008) on phenomena traditionally regarded as manifestations of modality (such as mood choice in the complement clause) leaves complementizers unattended. However, as I aim to show, the complementizer *hogy*, as a part of construction formed of a complementizer and additional morphosyntactic elements (mood markers or the interrogative clitic *-e*, as well as special omission patterns in an adequate semantic or pragmatic context), is able to express modal meanings.

I will argue below that the complementation strategies of Hungarian show a distribution based primarily on modal functions. Consider the different uses of *hogy*-constructions:

Type 1: declarative or neutral use: non-modal (neutral) matrix predicate + *hogy* + indicative;
Type 2a: modal[2] use: modal matrix predicate + *hogy* + non-indicative mood markers;
Type 2b: modal use: modal matrix predicate + (*hogy*) + interrogative clitic;
Type 2c: evidential use: hearsay / perception matrix predicate + (*hogy*) + indicative[3].

2 It has to be emphasized again that the terms modal and modality are used in a very broad sense here.

3 An interesting relation between Types 2a and 2c (indicating also the closeness of the categories 'epistemicity' and 'evidentiality') is the kind of evidential use of epistemic matrix verbs. If the verbs *hisz* 'believe' or *gondol* 'think' are used with epistemic meaning, they license the pattern *hogy* + conditional. As evidentials they trigger the structure (*hogy*) + indicative in the complement clause. (The different interpretation possibilities are also supported by the different omission patterns of the complementizer, see Section 3.) The sentence in (i) can be interpreted as both evidential (inferential) and epistemic:

(i) *Azt gondolom / nem gondolom* [*hogy Józsi elmegy a boltba*].
that think:0.1SG not think:0.1SG COMP Józsi PREF:go[IND.3SG] the shop:ILL
'I think / I don't think that Józsi will go to the shop.'

Conditional, as the example shows, is also acceptable in the complement clause:

(ii) *Azt gondolom / nem gondolom,* [*hogy Józsi elmenne a boltba*].
that think:0.1SG not think:0.1SG COMP Józsi PREF:go:COND[3SG] the shop:ILL
'I think / I don't think that Józsi will go to the shop.'

However, example (ii) (in which the negative matrix predicate signals uncertainty on part of the speaker) allows only epistemic interpretation. This claim is supported by two facts: first, modal operators are tolerable in such complement clauses while evidentials are not, witness (iii) and (iv); secondly, the epistemic matrix predicates (like epistemic matrix predicates generally) do not license complementizer omission (see Section 3). It should be also noted that an affirmative evidential matrix predicate and the conditional mood in the embedded clause is not ungrammatical but in this case the embedded clause has conditional meaning.

(iii) *Azt gondolom / nem gondolom,* [*hogy Józsi esetleg/ *nyilván elmenne*
that think:0.1SG not think:0.1SG COMP Józsi possibly obviously PREF:go:COND[3SG]
a boltba].
the shop:ILL
'I think / I don't think that Józsi will possibly go to the shop.'

(iv) *Azt hiszem / nem hiszem,* [*hogy talán / *nyilvánvalóan két különböző*
that think:0.1SG not think:0.1SG COMP maybe obviously two different
betegségem lenne].
disease:POSS.1SG be:COND[3SG]
'I think / I don't think I might have two kinds of disease.'

2.3.1 Type 1

The Type 1 use of the complementizer *hogy* is associated with non-epistemic or unmarked (neutral or declarative) meaning. In this case the finite verb of the complement clause is in the indicative mood, and no other morphosyntactic or lexical element referring to epistemicity can occur in the clause. The matrix predicate can be both affirmative and negated:

(9) a. *Józsi tudja,* [*hogy Anna olvas*].
Józsi know:O.3SG COMP Anna read[3SG]
'Józsi knows that Anna is reading.'
b. *Józsi nem tudja,* [*hogy Anna olvas*].
Józsi not know:O.3SG COMP Anna read[3SG]
'Józsi does not know that Anna is reading.'

The subjunctive marker, the conditional marker, and modal adverbs are not allowed in the complement clause in this type of use; see (9a) and (9b). An interrogative clitic is grammatically possible but it cannot indicate a modally neutral proposition; see (9c).

c. *Józsi (nem) tudja,* [*hogy Anna *olvasson*].
Józsi not know:O.3SG COMP Anna read:SBJV:3SG
Intended: 'Józsi knows / does not know that Anna is reading.'
d. *Józsi (nem) tudja,* [*hogy Anna *biztos olvas*].
Józsi not know:O.3SG COMP Anna surely read[3SG]
e. *Józsi (nem) tudja,* [*hogy Anna olvas=e*].
Józsi not know:O.3SG COMP Anna read[3SG]=Q
'Józsi knows / does not know if Anna is reading.'

2.3.2 Type 2

In Type 2 uses, a modal matrix predicate selects a combination of the complementizer itself and additional morphosyntactic elements in the complement clause, such as non-indicative mood markers or the interrogative clitic *-e*.

2.3.2.1 Type 2a: modal matrix predicate + *hogy* + non-indicative mood markers

Hogy + non-indicative mood combinations are usually triggered by non-neutral matrix predicates. The occurring non-indicative moods are the subjunctive, the conditional, and sometimes the third person imperative, which is formally identical with the subjunctive proper; see (10). The selection of the moods in the complement clause is a complex phenomenon I cannot discuss in detail here, but

evidence suggests that the modal nature of the matrix predicate has an influence on it.

(10) *Józsi szeretné,* [*hogy Anna énekeljen*].
Józsi would.like:O.3SG COMP Anna sing:SBJV:3SG
'Józsi would like Anna to sing.'

Traditional analyses (e.g. Tompa 1962; Pataki 1984) consider the indicative vs. non-indicative opposition in Hungarian complement clauses as a manifestation of the 'realis vs. irrealis' opposition. However, this association between form (morphological mood) and meaning ('realis' vs. 'irrealis), is a tendency rather than a rule. Affirmative realis matrix predicates usually trigger the combination *hogy* + indicative in the complement clause, but there are also examples where a non-realis matrix verb (e.g. *elképzel* 'imagine' or *álmodik* 'dream') selects this pattern, see (11), and examples where both the indicative and non-indicative moods are equally possible, see (12).

(11) *Azt álmodtam,* [*hogy Anna énekel*].
that dream:PST:1SG COMP Anna sing[3SG]
'I dreamt that Anna was singing.'

(12) *Nem hiszem,* [*hogy egy ilyen jó ember, mint János elveszi / elvegye / elvenné ezt a nőt*].
not believe:O.1SG COMP a such nice man like János marry[IND]:O.2SG marry:SBJV:O.3SG marry:COND.O.3SG this:ACC the woman:ACC
'I don't believe that such a nice man as János marries / should marry / would marry that woman.'
(Tóth 2008: 86)

The rationale behind the interchangeability of the indicative and the non-indicative moods is rather captured by the distinction 'non-modal vs. modal', than by 'realis vs. irrealis'. Accordingly, modal matrix predicates allow non-indicative moods in the complement clause introduced by *hogy*, and non-modal matrix predicates license only the indicative; cf. (13a) and (13b). (The correlation between matrix predicates and modal phenomena in the complement clause will be discussed in greater detail in Section 3 below.)

(13) a. *Józsi örül,* [*hogy Anna eszik*].
Józsi be glad[3SG] COMP Anna eat[IND]:REFL/MID.3SG
'Józsi is glad that Anna is eating.'
b. *Józsi örül,* [*hogy Anna *egyen / *enne*].
Józsi be glad[3SG] COMP Anna eat:SBJV:3SG eat:COND[3SG]

A volitional matrix predicate as in example (10) triggers the subjunctive in the complement clause, and both epistemic necessity and epistemic possibility allow non-indicative moods.

(14) a. *Szükségszerű,* [*hogy a futástól elfáradunk / elfáradjunk*].
necessary COMP the running:ABL get tired[IND]:3PL get.tired:SBJV:3PL
'It is inevitable that we get tired of running'.
b. *Elég valószínűtlen,* [*hogy az üléstől elfáradunk / elfáradnánk / elfáradjunk*].
quite unlikely COMP the sitting:ABL get tired[IND]:3PL get.tired:COND:3PL get tired:SBJV:3PL
'It's quite unlikely that we get tired of sitting'.

Negation of the epistemic matrix predicate (indicating the uncertainty of the speaker) has a positive effect on licensing non-indicative moods in the complement clause; see e.g. (12) above.

2.3.2.2 Type 2b: modal matrix predicate + *(hogy)* + the interrogative clitic *-e*

The interrogative clitic *-e*, shortly presented in Section 1.3, forms a complementizing construction with the neutral complementizer *hogy* in embedded yes/no-questions after negated matrix predicates. In this section, the semantic functions of the interrogative clitic in complementation will be investigated. I claim that rather than functioning as a marker of indirect questions, the construction 'modal matrix predicate + (*hogy*) + interrogative clitic' indicates uncertainty.

The basic function of the interrogative clitic *-e* is the coding of yes/no-questions, although such questions can be expressed without it as well. Morphologically and syntactically unmarked yes/no-questions differ from declarative sentences only prosodically; cf. (15).

(15) *Józsi szereti Annát?*
Józsi love:O.3SG Anna:ACC
'Does Józsi love Anna?'

If the interrogative clitic is present, the special prosodic pattern of the unmarked yes/no-questions disappears. The clitic occurs on the finite verb:

(16) *Józsi* *szereti=e* *Annát?*
Józsi love:O.3SG=Q Anna:ACC
'Does Józsi love Anna?'[4]

The interrogative clitic occurs almost exclusively in embedded finite complement clauses, mostly after negated matrix predicates. In examples (17)–(19) an interrogative interpretation of the clitic seems plausible only at the first sight: although all complement clauses are quasi yes/no-questions, they do not convey true illocutionary meaning. Instead of questioning, they indicate that the subject of the matrix predicate has incomplete information about the complement proposition.

(17) *Józsi* *nem* *tudja,* [*hogy* *szereti=e* *Annát*].
Józsi not know:O.3SG COMP love:O.3SG=Q Anna:ACC
'Józsi doesn't know whether he loves Anna.'

(18) *Megnézem,* [*hogy* *megfőtt=e* *már* *az* *ebéd*].
PREF:look:O.1SG COMP PREF:cook:PST[3SG]=Q already the lunch
'I check if the lunch is ready.'

(19) *Gizike* *Albitól* *kérdezi,* [*hogy* *van=e* *felvétel...*]
Gizike Albi:ABL ask:O.3SG COMP be[3SG]=Q enrollment
'Gizike asks Albi if there is enrollment...' (HNC)

The embedded interrogative clause is only grammatical with matrix predicates which indicate that the speaker[5] is not aware / not certain / not informed about the event or situation described in the complement clause. Therefore, matrix

4 Although the question shown in (16) is grammatical, many speakers find this type of structure artificial or "incomplete". The clitic is more natural in non-neutral, emotionally colored contexts with an explicit modal element.

(i) *Hát* *Józsi* *szereti=e* *Annát?*
MOD.PRT Józsi love:O.3SG=Q Anna:ACC
'Józsi loves Anna, doesn't he?'

Here, the modal nature of the question is attested by the fact that speakers tend to answer with *szerintem igen* 'as for me, yes' / *szereti hát* '~he loves her, for sure' rather than with a simple yes or no.

5 There are cases, however, where the subject referent of the matrix predicate is certain about the situation described in the complement clause but someone else (in contrast) is not:

(i) *Tudom,* *hogy* *Józsi* *járt=e* *Nepálban.*
know:O.1SG COMP Józsi go:PST[3SG]=Q Nepal:INE
'I know if Józsi has been in Nepal' (but e.g. you don't).

(ii) *Már* *megnéztük,* *hogy* *megfőtt=e* *az* *ebéd.*
already PREF:look:PST:O.1PL COMP PREF:cook:PST.3SG=Q the lunch
'We checked already if the lunch was ready' (but e.g. the guests are not informed).

predicates which cannot express the above epistemic meanings, cannot select *-e* in the complement clause:

(20) *Józsi sajnálja,* [*hogy szereti*=e Annát*].
Józsi regret:O.3SG COMP love:O.3SG=Q Anna:ACC
'Józsi regrets *if he loves Anna.'

2.3.2.3 Type 2c: evidential matrix predicate + *(hogy)* + indicative

If the matrix predicate expresses evidential meaning, beside the complementizer no additional morphosyntactic means contribute to this meaning in the complement clause. Example (21) contains a hearsay matrix verb, and the matrix predicates in (22) convey that the speaker has only indirect evidence for the proposition expressed by the complement clause[6].

(21) *Azt mondják,* [*hogy a Bakonyban már leesett a hó*].
that say:O.3PL COMP the Bakony:INE already PREF:fall:PST[3SG] the snow
'People say that snow has already fallen in the Bakony mountains.'

(22) *Úgy látom / nézem / tudom / ...,* [*hogy Józsi megette a csokimat*].
so see / look / know:O.1SG COMP Józsi PREF:eat:O.3SG the chocolate:POSS.1SG:ACC
'As I see / look at it / as far as I know (~ as I gather), Józsi ate up my chocolate. '

The constructions shown in (21) and (22) seem to be very similar to Type 1 use of the complementizer (see Section 2.3.1 above). As examples (23) and (24) show, neither non-indicative moods, nor the interrogative clitic is acceptable in such complement clauses.

It is not the uncertainty of the speaker but the uncertainty of another relevant person in the speech situation that explains the use of the interrogative clitic *-e* in these examples.

6 The status of the matrix predicates in (22) is controversial. Kiss (1987: 150–151) regards matrix predicates of the type *azt hiszem* 'I believe / in my opinion' and hearsay predicates as interjections rather than real matrix predicates because the structure and meaning of these expressions is not influenced by the use of the element *úgy* 'so'. However, the categorization of this type of predicates as modal tags is problematic because their internal structure can be changed easily: although the 1SG present indicative form of the verb is the most natural and frequent one, other persons, numbers, and tenses are grammatical, too.

(23) *Hallottam,* [*hogy Józsi elutazik /* **elutazzon*[7]
hear:PST:O.1SG COMP Józsi PREF:travel[IND]:REFL/MID.3SG PREF:travel:SBJV:3SG
Nepálba].
Nepal:ILL
Intended: 'As I have heard, Józsi will travel to Nepal.'

(24) *Úgy hírlik,* [*hogy Józsi járt*=e Nepálban*].
so rumor:REFL/MID.3SG COMP Józsi go:PST[3SG]=Q Nepal:INE
Intended: 'It is rumored that Józsi has been in Nepal.'

The difference between Type 1 and Type 2c uses can only be captured by the distinctive omission pattern of the complementizer *hogy* after evidential matrix predicates; see Section 4.

3 Distribution of the different uses of *hogy* 'that'

In this section I investigate the distribution of the different uses of *hogy*-constructions. I will test the influence of the semantics of the matrix verb on the selection of the *hogy*-constructions described above and how veridicality[8] and changes of veridicality affect it. In particular, I will test the function of relevant polarity operators in order to clarify how modal and evidential attitudes are represented in the structure of the complement clause. I will also examine the possible effects of the negation of matrix predicates (and of predicates with inherently negative semantics) on the complement clauses.

3.1 Semantic types of matrix predicates

As pointed out in numerous studies, the semantics of the embedding verb appears to be responsible in many languages for mood selection in the complement clause (e.g. Bolinger 1968; Terrell & Hooper 1974; Farkas 1992a, 1992b; Haverkate 2002). The lexico-semantic properties of matrix predicates seem to play a crucial role in mood choice also in Hungarian, although (as it will be shown below) they cannot be the only factor responsible for mood-licensing in complement clauses (cf. Tóth 2008). Before turning to the analysis of matrix predicate semantics and its effect

7 The conditional is grammatical here if the sentence has the meaning 'the subject has the intention to travel'.

8 As veridicality has to do with relativization of the truth, it can be regarded as a modality-related notion.

on the selection of *hogy*-constructions, it has to be emphasized that the goal of the present section is neither to give an explanation of mood choice in Hungarian *hogy*-complements, nor to provide an exhaustive analysis of Q-licensing in complement clauses, but only to investigate the selection of different *hogy*-constructions (including *hogy* + indicative mood / non-indicative mood constructions, as well as *hogy* + interrogative clitic structures).

By analyzing the distribution of complementizer constructions on a semantic basis, I follow Tóth (2008) and classify matrix predicates of complement clauses as follows:

Table 1: Semantic types of complement taking predicates

Semantic type	Examples
Factives	*örül* 'be glad', *szomorkodik* 'be sad'
Semifactives[9]	*sajnál* 'regret' *elfelejt* 'forget'
Epistemics	*hisz* 'believe' *gondol* 'think'
Assertives	*mond* 'say', *tagad* 'deny'
Fiction verbs	*elképzel* 'imagine', *álmodik* 'dream'
Directives	*utasít* 'instruct', *megilt* 'forbid'
Permissives	*megenged* 'permit' *megakadályoz* 'inhibit'
Purposives	*igyekszik* 'endeavor', *törekszik* 'strive'
Desideratives (volitives)	*szeretne* 'would like', *kíván* 'wish'

3.1.1 *Hogy* + indicative/non-indicative moods

According to Tóth (2008: 112), the mood choice in complement clauses which is based on the veridicality status of the matrix verbs (cf. Giannakidou 1998) yields a tripartite division of the above semantic types. The first group of matrix predicates comprises true factive and semifactive verbs (as typical veridicals), the second group epistemics, assertives, and fiction verbs (categorized as relative veridicals), whereas the third group contains directive, permissive, purposive, and desiderative verbs (nonveridicals). Tóth (2008) argues that the indicative mood is licensed in the complement clause, if the matrix predicate is veridical or relative veridical (see [25] and [30]–[32] below) but is not licensed under nonveridicals (see examples [26]–[29]). The group of relative veridical matrix predicates is ambivalent: epistemics, assertives and fiction verbs allow both the indicative and non-indicative moods.

9 The category of semifactives includes also perception verbs.

Negation, as an operation that can influence the veridicality of a proposition (changing a veridical to nonveridical or antiveridical) plays a significant role in mood selection in *hogy*-complements. However, the effect is not always the same. A negated or an inherently negative matrix predicate does not affect the mood choice in the group of factive and semifactive embedding predicates: in both semantic types the indicative has to be used in the *hogy*-complement clause; cf. (25a) and (25b).

(25) a. *Józsi (nem) örül, [hogy Anna megérkezett*].
Józsi not be glad[3SG] COMP Anna PREF:arrive[IND]:PST[3SG]
'Józsi is (not) glad that Anna arrived.'
b. *Józsi (nem) örül, [hogy Anna *megérkezzen /*
Józsi not be glad[3SG] COMP Anna PREF:arrive:SBJV:3SG
**megérkezett volna*].
PREF:arrive:PST [3SG] AUX:COND

Likewise, negation of the matrix predicate (or affirmation of a predicate with inherently negative semantics) does not affect the choice of the imperative and the subjunctive in clear nonveridicals, in which the selection of the indicative leads to ungrammatical constructions; cf.:

(26) a. *Utasítottam Józsit, [hogy vegye fel /*
instruct:PST:O.1SG Józsi:ACC COMP pick.up:IMP:O.3SG PREF
**felveszi Annát az állomásnál*].
PREF:pick.up[IND]:O.3SG Anna:ACC the station:ADE
'I instructed Józsi to pick up Anna at the station.'
b. *Megtiltottam Józsinak, [hogy felvegye /*
forbid:PST:O.1SG Józsi:DAT COMP PREF:pick.up:SBJV:O.3SG
**felveszi Annát az állomásnál*].
PREF:pick.up[IND]:O.3SG Anna:ACC the station:ADE
'I forbade Józsi to pick up Anna at the station.'

(27) a. *Józsi megengedi, [hogy játsszak / *játszom a*
Józsi permit:O.3SG COMP play:SBJV:1SG play[IND]:REFL/MID.1SG the
kutyájával].
dog:POSS.3SG:INS
'Józsi lets me play with his dog.'
b. *Józsi nem engedi meg, [hogy játsszak *játszom*
Józsi not permit:O.3SG PREF COMP play:SBJV.1SG play[IND]:REFL/MID.1SG
a kutyájával].
the dog:POSS.3SG:INS
'Józsi doesn't let me play with his dog.'

(28) *(Nem) igyekszem, [hogy Anna kiegyensúlyozott legyen*].
not strive:REFL/MID.1SG COMP Anna balanced be:SBJV:3SG
I (don't) strive that Anna be well balanced.

(29) *Józsi azt akarja, [hogy Anna leüljön / üljön*
Józsi that want:O.3SG COMP Anna PREF:sit.down:SBJV:3SG sit.down:IMP:3SG
*le / *leül*].
PREF sit.down[IND.3SG]
'Józsi wants that Anna take a seat.'

On the other hand, inherently negative or negated matrix predicates affect the choice of complement mood in the 'relative veridical group' of verbs, which allow both the indicative and the conditional mood. The choice of the conditional indicates that the probability of the proposition expressed by the complement clause is weaker than in complements with the indicative; cf. the examples in (30) with epistemic verbs, (31) with the assertive 'say', and (32) with fiction verbs.

(30) a. *Józsi azt / úgy hiszi , [hogy Anna megérkezett*].
Józsi that / so believe:O.3SG COMP Anna PREF:arrive[IND]:PST[3SG]
'Józsi believes (so) that Anna arrived.'
b. *Józsi kétli, [hogy Anna megérkezett / megérkezett*
Józsi doubt:O.3SG COMP Anna PREF:arrive[IND]:PST[3SG] PREF:arrive.PST[3SG]
volna].
AUX:COND
'Józsi doubts that Anna arrived.'

(31) a. *(Azt) mondtad, [hogy Anna az utcán sétál*].
that say:PST:O.2SG COMP Anna the street:SPE walk[IND.3SG]
'You said that Anna was walking on the street.'
b. *Nem mondtad, [hogy Anna az utcán sétál /*
not say:PST:O.2SG COMP Anna the street:SPE walk[IND.3SG]
sétálna].
walk:COND[3SG]
'You didn't tell that Anna was walking on the street.'

(32) a. *Elképzelem, [hogy tele van a pénztárcám*].
imagine:O.1SG COMP full be[IND.3SG] the purse:POSS.1SG
'I imagine that my purse is full.'
b. *Nem is álmodom, [hogy valami van / lenne*
not PRT dream:O.1SG COMP something be[IND.3SG] be:COND[3SG]
a pénztárcámban].
the purse:POSS.1SG:INE
'I don't even dream that there is something in my purse.'

The veridicality status of the above verb types can vary. Semifactive verbs of perception can also express nonveridical meanings when relativized by the particle *úgy* 'so', see (33), and some assertives can be used as directives, see (34). However, it has to be emphasized that the change in veridicality does not automatically affect the selection of mood in the complement clause; cf. (33) which allows only the indicative mood in the complement clause.

(33) *Úgy* *látom,* [*hogy* *Józsi* *az* *utcán* *sétál /*
so see:0.1SG COMP Józsi the street:SPE walk[IND.3SG]
**sétálna /* **sétáljon*].
walk:COND[3SG] walk:SBJV:3SG
'I see that Józsi is walking on the street.'

(34) *Mondtam* *Józsinak,* [*hogy* *sétáljon* *az* *utcán*].
say:PST:0.1SG Józsi:DAT COMP walk:IMP:3SG the street:SPE
'I told Józsi to walk on the street.'

Thus, a pure semantic analysis of the matrix predicate types does not provide a definitive solution for the choice of grammatical mood in *hogy*-complements. As evidence suggests, beside conventional semantic properties, propositional attitudes of the speaker play an important role in the selection of mood in *hogy*-complements. It was shown that predicates with clear nonveridical or veridical semantics license non-indicative and indicative moods in *hogy*-complements, respectively. In these cases, even polarity operators, such as the modal element *úgy* 'so' or negation have no influence on mood choice. Nevertheless, in the relative veridical group there is a valid distribution on the basis of negation: if the matrix predicate is negated (or has inherently negative semantics), both *hogy*+indicative and *hogy*+non-indicative (conditional) are possible in the complement clause, the latter emphasizing the weaker probability of the complement proposition (from the point of view of the speaker).

3.1.2 *Hogy*+interrogative clitic *-e*

Now we turn to the distribution of *hogy* + interrogative clitic constructions and test how the above semantic types of matrix predicates and their veridicality status influence the selection of the interrogative clitic in complement clauses.

As it was already shown in Section 2.3.2.2, *hogy* and the interrogative clitic are compatible only when the subject of the matrix clause (or another relevant referent in the situation) has incomplete information about the proposition expressed by the complement clause. This criterion is manifested in two ways: either the matrix predicate is a "real question predicate" that selects a question-like complement (see e.g. Ginzburg 1993: 268 and den Dikken & Giannakidou 2002: 47 about "semiquestion-complements"), or it is a verb which needs an "answer-like" complement where the speaker has no information about the proposition or the information s/he has is incomplete. In these cases the interrogative particle indicates the uncertainty of the speaker.

To the predicates of the first group belong e.g. the following verbs: *kérdez* 'ask', *megvizsgál* 'investigate', *(meg)tudakol* 'seek information', while *(meg)*

tud 'know, come to know', *rájön* 'know, come to know', *kitalál* 'reveal', *elfelejt* 'forget' can express, depending on context, that the subject of the complement taking predicate is uncertain about the proposition expressed by the complement clause. This uncertainty is more emphasized when the matrix verb has a negative meaning or is negated.

Although both groups can select *hogy* + interrogative clitic complements in Hungarian, there are considerable semantic and syntactic differences between them. Predicates of the first group always trigger the interrogative clitic, see (35), whereas matrix predicates of the second group are felicitous both with and without the interrogative clitic; see (36).

(35) *Józsi (azt) kérdezte,* [*hogy Anna elment=e /* **elment*
Józsi that ask:PST:O.3SG COMP Anna PREF:GO:PST[3SG]=Q PREF:go:PST[3SG]
vásárolni].
shop:INF
'Józsi asked if Anna went shopping.'

(36) *Józsi elfelejtette,* [*hogy Anna elment=e /* *elment*
Józsi PREF:forget:PST:O.3SG COMP Anna PREF:go:PST[3SG]=Q PREF:go:PST[3SG]
vásárolni].
shop:INF
'Józsi forgot that Anna went shopping.'

Complement taking predicates allowing both complement constructions – with and without the interrogative clitic – are most typically semifactive or assertive. Nevertheless, numerous semifactive or assertive predicates are not well-formed with an interrogative *hogy*-complement, and it seems plausible that it is not the semantic type of the matrix verb but rather the propositional attitude of the subject (in the sense of van der Auwera and Ammann 2008) that triggers the interrogative clitic in the complement clause.

Given that the veridicality of the matrix verb turned out to have influence on the mood choice in the complement clause, it is worth testing if it also plays a role in the selection of the interrogative clitic. As pointed out by den Dikken and Giannakidou (2002), there is a distinction between the syntactic behavior of nonveridical and veridical matrix predicates in Hungarian. Nonveridical verbs (e.g. *kérdez* 'ask') require a question-like complement clause (as shown in [35]) and always select the interrogative clitic. Veridical verbs (e.g. *tud* 'know'), on the other hand, do not always trigger the clitic. Instead, they follow a simple selection pattern: if there is a negative operator in the matrix predicate or the proposition is modally not neutral (see also footnote 5), even an assertive predicate can select the interrogative clitic. This means that the interrogative clitic appears not to be selected by the verb itself but (in line with the observations of Adger and Quer

2001) is triggered rather by an appropriate polarity operator, which can be of e.g. negative, question or modal character. The use of the interrogative clitic in *hogy*-complements supports thus the evidence gained from the investigation of *hogy* + indicative/non-indicative mood constructions that in addition to the semantics of matrix predicates, the propositional attitude of the speaker is crucial in the selection of additional morphosyntactic elements in *hogy*-complement clauses.

3.1.3 (*Hogy*)+indicative

The matrix predicates selecting the (*hogy*)+indicative-construction in complement clauses (Type 2c) can all be considered to be verbs expressing that the speaker has only indirect evidence for the contents of the complement. Beside perception verbs and typical hearsay predicates, special uses of the verbs *mond* 'say' and epistemics such as *hisz* 'believe', *gondol* 'think' and *tud* 'know' can appear with this construction. The following distinction can be made between hearsay predicates and the other types of verbs: while 'hearsay' represents common knowledge and refers clearly to the source of evidence the speaker has about certain situation, perception verbs and the other verbs listed above are always more personal, and thus more sensitive to relativization of experience. Interestingly, perception verbs and epistemic modals, which otherwise behave differently as complement taking predicates, show similar behavior with respect to this construction: both types select the (*hogy*)+indicative in complement clauses, irrespective of the actual semantic status of the matrix predicate; cf. (37a) and (37b)[10].

(37) a. *(Úgy) érzem,* [*hogy Józsi már elkezdte pucolni*
so feel:O.1SG COMP Józsi already PREF:begin[IND]:PST:O.3SG clean:INF
a halat a konyhában].
the fish:ACC the kitchen:INE
'As I smell, Józsi has already begun with the cleaning of the fish in the kitchen.'

b. *Azt / úgy mondják,* [*sok hal van a Balatonban*].
that so say:O.3PL plenty fish be[IND.3SG] the Balaton:INE
'(As) people say, there is plenty of fish in Balaton.'

Concerning the veridicality status of the above types, perception verbs, epistemic and hearsay verbs can be interpreted either as veridicals or as nonveridicals, depending on the intention of the speaker. An epistemic reading of a verb which emphasizes the uncertainty of the speaker (judging the probability of the comple-

10 The meanings of the matrix predicates with the element *úgy* 'so' are slightly different from the predicates without. Native speakers of Hungarian estimate both versions as quasi equivalent in their meanings; however, some of them consider the contents of the *úgy*-type as less certain.

ment proposition) can be accepted as nonveridical but evidential readings are rather veridical. The ambiguity between evidentiality and epistemicity can be solved by a negative operator: the construction 'verb with an evidential use + (*hogy*) + indicative' (Type 2c) is not available if the matrix predicate is negated: both (38a) and (38b) employ other *hogy*-constructions typical for relative veridical verbs, namely either Type 1 '*hogy*+indicative' (indicating a neutral attitude of the speaker toward the comment proposition) or Type 2a '*hogy*+conditional' (expressing that the probability of the proposition expressed by the comment clause is relatively low).

(38) a. *Nem* *hallottam,* [*hogy* *Anna* *tegnap* *megérkezett /*
not hear:pst:O.1SG COMP Anna yesterday PREF:arrive:PST[3SG]
megérkezett *volna*].
arrive:PST[3SG] AUX:COND
'I didn't hear that/whether Anna has arrived yesterday.'

b. *Nem* *mondják,* [*hogy* *Józsi* *lánya* *szép /* *szép*
not say:O.3PL COMP Józsi daughter:POSS.3SG nice[IND.3sg] nice
lenne].
be:COND[3SG]
'They (?people) don't say that/whether Józsi's daughter is beautiful.'

These examples show that the evidential use seems to be compatible only with positive, veridical interpretations of the complement proposition. Thus, the 'evidential matrix verb + (*hogy*) + indicative mood' construction represents a use which is characteristic for a very small group of verbs and is limited to assertive predicates.

3.2 Summary

As shown above, the distribution of the four types of *hogy*-constructions depends on distinct semantic and pragmatic factors. In addition to the veridicality of the matrix predicate, the propositional attitude of the speaker influences the selection of different types of *hogy*-complement clauses.

The choice for *hogy*+non-indicative moods turned out to be influenced primarily by semantic factors, namely the semantic type and (respectively) the veridicality of the matrix predicate (cf. Tóth 2008). Accordingly, veridical matrix predicates select the indicative mood, while nonveridicals select non-indicative moods. In the relative veridical group of verbs the choice between the indicative and the conditional is decided on semantic and pragmatic grounds: if the matrix predicate is negated or has modal semantics (changing veridical meaning to non- or antiveridical), the conditional is selected. Both negation and modality can be

also considered as propositional attitudes of the speaker indicating low degree of certainty about the proposition expressed by the complement clause.

Hogy + interrogative clitic constructions also form two groups that differ in terms of veridicality. Nonveridical predicates expressing questions always select the interrogative clitic in the complement clause, but veridicals select this clitic only if the matrix predicate occurs with modally non-neutral complement or is negated. In the latter case the *hogy* + interrogative clitic construction expresses the uncertainty of the speaker.

The use of the type (*hogy*) + indicative is chosen in connection with perception or hearsay predicates where, unlike in the above cases, an epistemically non-neutral matrix predicate does not trigger the conditional mood. Negated predicates do not license this construction.

4 Complementizer omission

This section focuses on the special omission patterns of the complementizer *hogy* in the construction Types 1–2c.

4.1 The possibility for complementizer omission

As shown in Section 1.3, there are various types of simple and complex complementizers in Hungarian. However, there are no zero complementizers: complement clauses without overt complementizers can always be considered as cases of complementizer omission.

According to Kenesei (1994a: 4, following Stowell 1981), complementizers are more easily deletable if they are recoverable. Recoverability can be ensured in various ways, e.g. by wh-phrases, complementizer combinations where a part of the complex complementizer is not omittable, or additional morphosyntactic features in the complement clause. As cross-linguistic evidence suggests, the modality of the predicate of the complement clause (see Kenesei & de Urbina 1994: 4, referring to Vamling 1989 and Ransom 1986 for Georgian) and the mood choice in the complement clause (e.g. the subjunctive in Spanish) can have an influence on complementizer omission. Hungarian data, as it will be shown below, partly support this observation.

In sentences with a wh-element in the complement clause, the neutral complementizer *hogy* (as an element without lexical meaning) can be omitted without any loss of information:

(39) *Józsi tudja,* [*[hogy] miért / mikor, etc.* *megy* *Anna* *moziba*].
Józsi know:0.3SG COMP why when go[3SG] Anna cinema:ILL
'Józsi knows why/when etc. Anna goes to the cinema.'

Likewise, if there is a complex complementizer formed with *hogy*, the lexically neutral part can be omitted:

(40) *Józsi örül,* [*[hogy]ha* *Anna* *moziba* *megy*].
Józsi be glad[3SG] COMP Anna cinema:ILL go[3SG]
'Józsi is glad if Anna goes to the cinema.'

However, in some sentences where *hogy* is the only (full) complementizer, recoverability turns out to be insufficiently ensured to allow the deletion of *hogy*.

4.2 Omission of *hogy* in the constructions Type 1–2c

4.2.1 Type 1: neutral use (non-modal matrix predicate + *hogy* + indicative)

With *hogy* as the only (not complex and not combined) complementizer in the sentence, the omission of the complementizer is not possible. There is no other morphosyntactic element that could identify the complement clause as a complement in (41) and therefore *hogy* is not omissible.

(41) *Józsi olvasta,* [*hogy* **[hogy]* *a* *medve* *barna*].
Józsi read:PST:0.3SG COMP the bear brown.
'Józsi read that bears are brown.'

4.2.2 Type 2a: modal matrix verb + *hogy* + non-indicative moods

Although in Type 2a the complementizer is combined with the subjunctive, the conditional or the imperative, mood marking alone does not prove enough to identify the complement as such. Type 2a constructions require that the complementizer is overt; cf. the examples with different verbs in (42).

(42) a. *Azt akarom,* [*hogy* **[hogy]* *eperfagyit* *hozzál*].
that want:0.3SG COMP strawberry:ice.cream:ACC bring:SBJV:2SG
'I want you to bring strawberry ice cream.'

b. *Nem tartom valószínűnek, [hogy *[hogy] Anna tegnap*
not find:O.1SG probable:DAT COMP Anna yesterday
megérkezett volna].
PREF:arrive:PST[3SG] AUX:COND
'I don't find it probable that Anna has arrived yesterday.'

c. *A tábornok megparancsolta [a katonának, hogy *[hogy] menjen fel a toronyba*].
the general PREF:command:PST:O.3SG the soldier:DAT COMP go:IMP:3SG PREF the tower:ILL
'The general ordered the soldier to go up to the tower.'

These examples show that the use of non-indicative moods in the complement clause is not grammatical without the complementizer *hogy*.

4.2.3 Type 2b: non-neutral matrix verb + (*hogy*) + interrogative clitic

It was already mentioned in Section 1.3 that the interrogative clitic *-e* forms a complementizer construction with the neutral complementizer *hogy*. This construction functions as an embedded question or has modal functions. The interrogative clitic tends to be more independent in its complementizer function than are the grammatical moods discussed above in Section 4.2.2. This is reflected in the omission pattern of Type 2b constructions. Concerning the omissibility of *hogy*, the distinction between the two groups of Type 2b constructions (recall Section 3.1.2) seems to be reasonable: the complementizer *hogy* introducing question-like complements (selected by matrix verbs like *kérdez* 'ask') is more felicitous with overt complementizer in the complement clause[11]; in the case of embedded interrogatives indicating uncertainty on the part of the speaker, the complementizer *hogy* is optional. Compare (43a) and (43b)

(43) a. *Józsi felhívta Annát, [hogy ?[hogy] elkészült=e]*
Józsi PREF:call:PST.O.3SG Anna.ACC COMP PREF:get.ready:PST[3SG]=Q
'Józsi called Anna on the phone (to ask) if she got ready.'

b. *Józsi nem tudja biztosan, [[hogy] Anna meghívta=e].*
Józsi not know:O.3SG surely COMP Anna invite:PST:O.3SG=Q
'Józsi doesn't know for sure whether Anna has invited him.'

Although both these types contain the interrogative clitic as a possible trace to ensure recoverability, it tends to function as such better in epistemic contexts.

11 However, the most frequent verb from this group *kérdez* 'ask' allows complementizer omission. Other verbs (e.g. *megtudakol* 'ask, ask for information', *érdeklődik* 'ask, ask for information') sound better with an overt complementizer.

4.2.4 Type 2c: evidential matrix verb + (*hogy*) + indicative

Type 2c uses show that complementizer omission is allowed in evidential contexts, regardless of additional interrogative or modal elements selected in the complement clause. As examples (44a) and (44b) show, the omission is grammatical in sentences with a perception or hearsay matrix predicate.

(44) a. *Látom,* [*[hogy]* *Anna* *már* *elutazott*].
see:O.1SG COMP Anna already PREF:leave:PST[3SG]
'I see that Anna has left already.'
b. *Azt* *mondják,* [*[hogy]* *Józsi* *lánya* *szép*].
that say:O.3PL COMP Józsi daughter:POSS.3SG beautiful
'People say that Józsi's daughter is beautiful.'

It was pointed out in Section 3.1.3 that a negated matrix predicate does not allow for evidential interpretation of the construction. Accordingly, complementizer omission is not possible in examples (45a) and (45b).

(45) a. *Nem* *hallom,* [**[hogy]* *a* *szomszéd* *otthon* *van*].
not hear:O.1SG COMP the neighbor home be[3SG]
Intended: 'I don't hear that the neighbor is at home.'
b. *Nem* *beszélik,* [**[hogy]* *Anna* *fia* *megszületett*].
not talk:O.3PL COMP Anna son:POSS.3SG be.born:PST[3SG]
Intended: 'People don't say that Anna's son has been born.'

4.3 Summary

Omission patterns of the different *hogy*-constructions show that there is a connection between the omissibility of *hogy* and modality. As it seems, additional morphosyntactic elements (mood markers and the interrogative clitic) do not ensure the syntactic recoverability of the deleted complementizer: *hogy* is not always omissible in these constructions. Omission is sensitive rather to modal features: it is acceptable in contexts of uncertainty and evidentiality[12] but not possible in modally neutral contexts or in modal contexts expressing probability or volition in which the complement verb is in one of the non-indicative moods.

12 The relationship between epistemic necessity and evidentiality is well researched in the literature (e.g. van der Auwera & Plungian 1998; de Haan 1999; Boye 2012, for Hungarian Kiefer 2005; Bartos 2008).

5 Non-complementizing functions of *hogy*

The functions of the complementizer *hogy* are not restricted to complementation. Similarly to many particle-like elements, it can be used also as an emphasizer, giving the utterance strong emotive characteristics. In this function it can occur optionally also in root wh-exclamatives (see Lipták 2005: 164):

(46) *Hogy* *Józsinak* *mekkora* *mázlija* *van!*
COMP Józsi:DAT how.big good.luck:POSS.3SG be[3SG]
'How lucky Józsi is!'

The utterance in (46) could be also interpreted as a fracture of the following sentence, including *hogy* in its genuine complementizer function:

(47) *Elképesztő/hihetetlen/őrület,* [*hogy* *Józsinak* *mekkora* *mázlija* *van*]*!*
incredible/unbelievable/awesome COMP Józsi:DAT how.big good.luck:POSS.3SG be[3SG]
'It's incredible/unbelievable/awesome how lucky Józsi is!'

Another function of *hogy* to be mentioned is its use as an interrogative particle and a proadverb with the meaning 'how'[13]. In the latter case, *hogy* has also a longer form, *hogyan* 'how':

(48) *Hogy* *fogod* *letenni* *ezt* *a* *nehéz* *vizsgát?*
how AUX:O.2SG PREF:put:INF this the hard exam:ACC
'How will you pass this hard exam?'

(49) *Józsi* *hallgatta,* *(hogy)* *hogy(an)* *esik* *az* *eső.*
Józsi listen:PST:O.3SG COMP how fall:REFL/MID.3SG the rain
'Józsi listened how the rain was falling.'

Example (49) shows that the complementizer *hogy* can be immediately followed by the element *hogy* 'how', which fills the focus position in the clause. Accordingly, unlike the complementizer *hogy*, this can bear prosodic emphasis.

---

13 *Hogy* 'how' has also a special, slightly archaic (and nonstandard) questioning function typical for price requests on the market place:

(viii) *Hogy* *a* *paprika?*
how the paprika
'What does the paprika cost?'

Another, clearly archaic use of *hogy* as an adverb or as a conjunction '~that way, as, in that manner' (usually together with *úgy* 'so, in that manner'), is very rare, but still intelligible. The relative adverbial *ahogy* '~that way, as' is used in this function in contemporary standard Hungarian.

6 The grammaticalization of *hogy*

The complementizer *hogy* is one of the oldest subordinators in Hungarian. It occurs in this function already in Old Hungarian and can be analyzed as a compound of the pronominal stem *ho-* (originating in the Uralic **ku-*/**ko-* pronominal stem, identical with the stems of many pronouns and adverbials in the contemporary Uralic languages[14]) and the locative *-l* palatalized and fricativized to *-gy*. Its primary function has probably been adverbial. This function is preserved in the contemporary manner adverbial function (in the form *(a)hogy* 'in that manner, like'). The manner adverb gained also interrogative and relative functions (as shown in Section 5), which paved the way for the rise of the paired coordinating conjunction *úgy – (a)hogy* 'as – as'. A different grammaticalization cline can be established for the complementizer *hogy*, which most probably developed directly from the original adverb *hogy* 'so, in that manner' and/or through a modal particle stage (cf. the exclamative use of *hogy* in Section 5); for details on the history and the functional change of *hogy* see Juhász (1991, 1992) and Haader (1991, 1995). Therefore, although both grammaticalization paths lead to conjunctions, it is the path leading to a complementizer that passes through the modal domain. Figure 1 shows the arguable grammaticalization cline(s) from manner adverb to coordinating conjunction and to complementizer. It has to be emphasized that all functions included in the figure are preserved in the contemporary language.

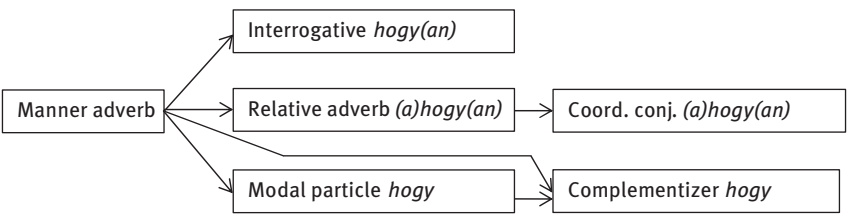


Figure 1. Grammaticalization of the complementizer *hogy*

14 For example, the Hungarian *hol* 'where (locative)', *hova* 'where (lative)', *honnan* 'where from'; the Mansi *kan* 'who', *kāt* 'where'; the Finnish *kuka* 'who', *kuin* 'than, as', the Tundra Nenets *xíbya* 'who', *xurka* 'what kind', etc.

7 Summary

The basic aim of this paper was to investigate whether the complementizer *hogy* 'that' can be considered as a modal element in Frajzyngier's sense or should be categorized as a neutral subordinator with only syntactic functions. Given that this complementizer often co-occurs with certain morphosyntactic elements (such as mood markers and the interrogative clitic *-e*) in the dependent clause, forming with them construction types specific for complementation, it proved to be reasonable to analyze it as a part of construction. It was shown that the construction-specific complementizer *hogy* exhibits distinctive behavior, which depends on the semantics of the matrix predicate and the propositional attitude of the speaker. Thus, different *hogy*-constructions can express modally neutral meanings, on the one hand, and epistemic meanings as well as evidentiality, on the other. *Hogy*, selected by a veridical matrix predicate, combined with the indicative mood in the comment clause indicates modally neutral statements. Nonveridical semantics of the matrix verb triggers the combinations *hogy*+imperative and *hogy*+subjunctive in the complement clause. Matrix predicates which allow both veridical and nonveridical interpretations can select *hogy*+conditional in the complement clause, if the degree of certainty about the truth of the complement proposition is low. The construction formed by *hogy* and the interrogative clitic *-e* shows a dual behavior: it can indicate that the subject of the matrix clause has incomplete information about the proposition expressed by the complement clause (a question-like use), or it can express uncertainty on the part of the subject (a modal use). Both uses allow for complementizer omission. Finally, a small group of verbs with evidential semantics selects the (*hogy*)+indicative-construction which seems to undermine the traditional treatment of the complementizer *hogy* as semantically neutral.

Sections 5 and 6 showed that *hogy* can be considered as a polyfunctional element in contemporary Hungarian and that all its functions can be derived from the function of manner adverb.

References

Adger, David & Joseph Quer. 2001. The syntax and semantics of unselected embedded questions. *Language* 77. 107–133.

van der Auwera, Johan & Vladimir Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2. 79–124.

van der Auwera, Johan & Andreas Ammann. 2008. Situational possibility. Epistemic possibility. Overlap between epistemic and situational possibility. In Martin Haspelmath, Matthew

Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures online*, Chapters 74, 75 and 76. Munich: Max Planck Digital Library. http://wals.info/feature/74, http://wals.info/feature/75, http://wals.info/feature/76 (accessed July 2014).

Bácskai-Atkári, Júlia. 2011. A komparatív operátor esete a mondatbevezetővel: Szintaktikai változások a magyar hasonlító mellékmondatokban [The comparative operator's affair with the complementizer: Syntactic changes in Hungarian subclauses]. In Katalin É. Kiss & Attila Hegedűs (eds.), *Nyelvelmélet és diakrónia*, 103–119. Budapest & Piliscsaba: Szent István Társulat.

Bácskai-Atkári, Júlia. 2012. *The history of Hungarian complex complementizers*, Manuscript, http://www.nytud.hu/oszt/elmnyelv/mgtsz/2012/108_b_a_j_munchen_manuscript.pdf (accessed July 2014).

Bácskai-Atkári, Júlia. 2013. Reanalaysis in Hungarian comparative subclauses. In Johan Brandtler, Valéria Molnár & Christer Platzack (eds.), *Approaches to Hungarian*, vol. 13: *Papers from the 2011 Lund conference*, 5–32. Amsterdam & Philadelphia: Benjamins.

Bartos, Huba. 2008. Mód és modalitás a magyar mondattanban [Mode and Modality in Hungarian Syntax]. *Nyelvtudományi Közlemények* 105. 80–103.

Bhatt, Rakesh & James Yoon. 1992. On the composition of COMP and parameters of V2. In Dawn Bates (ed.), *The proceedings of the Tenth West Coast Conference on Formal Linguistics*, 41–52. Stanford, CA: Center for the Study of Language and Information.

Bolinger, Dwight. 1968. *Aspects of language*. New York: Harcourt, Brace, and World.

Boye, Kasper. 2012. *Epistemic Meaning. A crosslinguistic and functional-cognitive study* (Empirical Approaches to Language Typology 43). Berlin: Mouton de Gruyter.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective* (Oxford Studies in Comparative Syntax). Oxford: Oxford University Press.

Condoravdi, Cleo. 2002. Temporal interpretation of modals: Modals for the Present and for the Past. In David I. Beaver, Stefan Kaufmann, Brady Z. Clark & Luis D. Casillas Martínez (eds.), *The construction of meaning*, 59–88. Stanford, CA: Center for the Study of Language and Information.

den Dikken, Marcel & Anastasia Giannakidou. 2002. From *hell* to polarity: "Aggressively non-D-linked" *wh*-phrases as polarity items. *Linguistic Inquiry* 33. 31–61.

Farkas, Donka. 1992a. Mood choice in complement clauses. In István Kenesei & Csaba Pléh (eds.), *Approaches to Hungarian,* vol. 4, 207–225. Szeged: JATE.

Farkas, Donka. 1992b. On the semantics of subjunctive complements. In Paul Hirschbühler & Konrad Koerner (eds.), *Romance languages and modern linguistic theory*, 69–104. Amsterdam: John Benjamins.

Frajzyngier, Zygmunt & Robert Jasperson. 1991. That-clauses and other complements. *Lingua* 83. 133–153.

Frajzyngier, Zygmunt. 1995. A functional theory of complementizers. In Joan L. Bybee & Suzanne Fleischman (eds.), *Modality in grammar and discourse*, 473–502. Amsterdam: John Benjamins.

Galambos, Dezső. 1907. *Tanulmányok a magyar relatívum mondattanáról* [Studies on the syntax of Hungarian relatives]. Budapest: Athenaeum.

Giannakidou, Anastasia. 1998. *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins.

Ginzburg, Jonathan. 1993. Propositional and non-propositional attitudes. In Peter Aczel, David Israel, Stanley Peters & Yasuhiro Katagiri (eds.), *Situation Theory and its applications,* vol.

3 (Center for the Study of Language and Information – Lecture Notes [Book 37]), 265–302. Stanford, CA: Center for the Study of Language and Information.
Haader, Lea. 1991. Az alárendelő mondatok: Az alanyi, állítmányi, tárgyi és határozói mellékmondatok [Subordinate clauses: Subjective, predicative, objective and adverbial subclauses]. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana I.: A korai ómagyar kor és előzményei*, 728–741. Budapest: Akadémiai Kiadó.
Haader, Lea. 1995. Az alárendelő mondatok: Az alanyi, állítmányi, tárgyi és határozói mellékmondatok [Subordinate clauses: Subjective, predicative, objective and adverbial subclauses]. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana,* vol. II/2: *A kései ómagyar kor: Mondattan. Szöveggrammatika*, 506–665. Budapest: Akadémiai Kiadó.
de Haan, Ferdinand. 1999. Evidentiality and epistemic modality: Setting boundaries. *Southwest Journal of Linguistics* 18(1). 83–101.
Haverkate, Henk. 2002. *The syntax, semantics and pragmatics of Spanish mood*. Amsterdam: John Benjamins.
Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language.* Cambridge: Cambridge University Press.
Juhász, Dezső. 1991. A kötőszók [Conjunctions]. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana,* vol. I: *A korai ómagyar kor és előzményei*, 476–500. Budapest: Akadémiai Kiadó.
Juhász, Dezső. 1992. A kötőszók [Conjunctions]. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana*, vol. II/2: *A kései ómagyar kor: Morfematika*, 772–814. Budapest: Akadémiai Kiadó.
Kálmán, László (ed.). 2001. *Magyar leíró nyelvtan – Mondattan I* [Hungarian descriptive grammar. Syntax]. Budapest: Tinta Kiadó.
Karttunen, Lauri. 1972. Possible and Must. In John P. Kimball (ed.), *Syntax and semantics,* vol. 1, 1–20. New York: Academic Press.
Kenesei, István, Robert M. Vago & Anna Fenyvesi. 1998. *Hungarian*. London & New York: Taylor & Francis Routledge.
Kenesei, István. 1992. On Hungarian complementizers. In István Kenesei & Csaba Pléh (eds.), *Approaches to Hungarian: The structure of Hungarian*, vol. 4, 37–50. Szeged: JATE.
Kenesei, István. 1994a. Complementation in Finno-Ugric, Manuscript, http://www.nytud.hu/kenesei/publ/finnugr2.pdf (accessed June 2014).
Kenesei, István. 1994b. Subordinate clauses. In Ferenc Kiefer & Katalin É. Kiss (eds.), *The syntactic structure of Hungarian: syntax and semantics,* vol. 27, 275–354. San Diego & New York: Academic Press.
Kenesi, István & Jon Ortiz de Urbina. 1994. Functional categories in complementation, Manuscript, http://www.nytud.hu/kenesei/publ/funccat.pdf (accessed July 2014).
Kiefer, Ferenc. 1981. What is possible in Hungarian? *Acta Linguistica Hung* 31. 147–185.
Kiefer, Ferenc. 1985. A *-hat/-het* képző jelentéstanához. Az episztemikus *-hat/-het* [On the semantics of the derivational suffix *-hat/-het*. The epistemic *-hat/-het*]. *Általános Nyelvészeti Tanulmányok* 16. 131–153.
Kiefer, Ferenc. 2005. *Lehetőség és szükségszerűség: Tanulmányok a nyelvi modalitás köréből* [Possibility and necessity. Studies on modality in Hungarian]. Budapest: Tinta Könyvkiadó.
Kirsner, Robert S. & Sandra A. Thompson. 1976. The role of pragmatic inference in semantics: A study of sensory verb complements in English. *Glossa* 10. 200–240.
Kiss, Katalin É. 1987. *Configurationality in Hungarian*. Dodrecht: Reidel.
Kiss, Katalin É. 2002. *The syntax of Hungarian*. Cambridge: Cambridge University Press.

Kratzer, Angelika. 1981. The notional category of modality. In Hannes Reiser & Hans J. Eikmeyer (eds.), *Words, worlds, and contexts*, 38–74. Berlin: De Gruyter.

Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–650. Berlin: De Gryuter.

Lipták, Anikó. 2005. The left periphery of Hungarian exclamatives. In Laura Brugè, Giuliana Giusti, Nicola Munaro, Walter Schweikert & Giuseppina Turano (eds.), *Contributions to the thirtieth Incontro di Grammatica Generativa*, 161–183. Venezie: Cafoscarina.

Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 2: *Complex constructions*, 2nd edn., 52–150. Cambridge: Cambridge University Press.

Nordström, Jackie. 2010. *Modality and subordinators* (Studies in Languages, Companion Series 116). Amsterdam: John Benjamins.

Pataki, Pál. 1984. A francia subjonctif és a magyar kötőmód [The French *subjonctif* and the Hungarian Subjunctive]. *Általános Nelvészeti tanulmányok* 15. 208–218.

Portner, Paul. 1997. The semantics of mood, complementation and conversational force. *Natural Language Semantics* 5.167–212.

Ransom, Evelyn. 1986. *Complementation: its meanings and forms*. Amsterdam: John Benjamins.

Sass, Bálint. 2009. "Mazsola" – eszköz a magyar igék bővítményszerkezetének vizsgálatára [Mazsola – A verb argument browser for Hungarian]. In Tamás Váradi (ed.), *Válogatás az I. Alkalmazott Nyelvészeti Doktorandusz Konferencia előadásaiból*, 117–129. Budapest: MTA Nyelvtudományi Intézet.

Stowell, Tim. 1981. *Origins of phrase structure*. Cambridge, MA: MIT dissertation.

Szabolcsi, Anna. 1992. Subordination: articles and complementizers. In István Kenesei & Csaba Pléh (eds.), *Approaches to Hungarian 4: The structure of Hungarian*, 123–140. Szeged: JATE.

Terrell, Tracy & Joan Hooper. 1974. A semantically based analysis of mood in Spanish. *Hispania* 57. 484–494.

Tompa, József (ed.). 1962. *A mai Magyar nyelv rendszere* [The system of contemporary Hungarian. Syntax]. Budapest: Akadémiai Kiadó.

Tóth, Enikő. 2006. Az alárendelt mellékmondatbeli módválasztásról [Mood choice in complement clauses]. In Balázs Sinkovics (ed.), *Lingdok 5*, 181–200. Szeged: JATE Press.

Tóth, Enikő. 2008. *Mood choice in complement clauses: a semantic approach with special reference to Hungarian* (Metalinguistica 21). Frankfurt am Main: Peter Lang.

Vamling, Karina. 1989. *Complementation in Georgian*. Lund: Lund University Press.

Váradi, Tamás. 2002. The Hungarian National Corpus. In *Proceedings of the 3rd International Conference on Language Resources and. Evaluation (LREC2002)*, 385–389. Las Palmas de Gran Canaria: European Language Resources Association.

Birsel Karakoç and Annette Herkenrath

Clausal complementation in Turkish and Noghay in a semantic perspective

1 Introduction

The topic of the present paper is linguistic means for embedding complement clauses, which we cautiously refer to as "complementation markers", in two Turkic languages: Turkish and Noghay.[1] The aim is to describe the forms involved and to discuss their functions. We briefly define complementation as the embedding of a clause – i.e. a functional and syntactic unit realizing a proposition or a predication and projecting an inner argument structure within itself – into the argument structure of a superordinate clause.

Typologically, the most central features of complementation, and of subordination more generally, in Turkic languages are non-finite, head-final ("left-branching") structures involving bound morphemes. These morphemes arguably fulfil functions that to a certain extent are comparable to those of free subjunctors[2] such as *that* in English. We argue that an analysis of complementation in languages of the Turkic type has to centrally include – in fact, must focus on – these bound morphemes as linguistic means of subordination, some of which can also occur in other syntactic functions that are not necessarily subordinating. The complementizer status of the said morphemes is a controversial issue.

This paper looks at non-finite complement clauses exclusively, leaving aside some rather controversial finite constructions.[3] Example (1) illustrates the kind of construction we have in mind:

1 In this article, we follow Turcological conventions in denoting languages, such as "Noghay" and "Azerbaijanian" (see Johanson & Csató [1998] 2006), although we are aware of the slightly different denotation in Ethnologue, which is otherwise followed in this volume, and which would be "Nogai" and "Azerbaijani", respectively.

2 We use "subjunctor" in the sense of Johanson (e.g. 2013).

3 Turkish and Noghay do have finite constructions that are sometimes analysed as subordinate, marked by *ki* (originally from Persian) and *diye* in Turkish, and *dep* in Noghay (converbial forms of the *verbum dicendi de-* 'say'), as well as procedures based on syntactic adjacency. Some approaches analyse Turkish *ki* in particular as an Indo-European-type complementizer (Kornfilt 1997: 46, 2007: 309–310 for a problematization; Göksel & Kerslake 2005: 409; Kerslake 2007; Coşkun 2010a, 2010b, 2012). Nonetheless, its status is rather controversial (see Johanson 1975a, 1990b; Rehbein 2006), and we follow an approach that does not consider *ki* as a complementizer. The case of *diye/dep* is different, such that it seems to be a better candidate for a complementizer – if one leaves aside the converbial inner structure of these elements and also leaves aside

Turkish

(1) [*Kaybol-duğ-un-u*] *ve* *ad-ın-ı* *ve* *soyad-ın-ı* *söyle*.
be_lost-DIK-POSS.2SG-ACC and name-POSS.2SG-ACC and surname-POSS.2SG-ACC say
'Say that you are lost and your name and surname.'
(Rehbein 2009: EFE08tk_GOZ_m_0402_ENF)[4]

The short complement clause in example (1) consists of a non-finite verbal predication *kaybol-duğ-un-u* [be_lost-DIK-POSS.2SG-ACC] 'that you are lost', subordinated to the transitive matrix verb *söyle-* 'say', which governs it as an accusative-marked direct object. The higher predicate is placed in the final position, after the complement clause. Note the parallel structure in the three objects to the matrix predicate: one clausal: *kaybol-duğ-un-u* 'that you are lost', and two lexical nouns: *ad-ın-ı* 'your name' and *soyad-ın-ı* 'your surname', all three carrying the same possessive and case suffixes.

Our paper focuses on the semantics of the verbal morphemes, such as *-DIK* in example (1), used in non-finite complement structures. We ask questions pertaining to their functional semantics at three levels: typological, inter-lingual, and intra-lingual. By "semantics" we mean the relationship between a specific form (in some theoretical approaches a "surface structure") on the one hand and an underlying and less language-specific functional concept on the other. This question can be reframed as pertaining to the functional microstructure of the specific forms in question, or to the functional quality inherent in their form. These relations are, as will be seen, typologically diverse.

This leads to the first level of our description, the level of typological comparisons. The syntax-theoretical concept of the "complementizer", when looked at with an Indo-European typology in mind, may simply and unproblematically denote a word unit such as English *that*, a morphologically unbound, construction-initially surfacing expression[5] with a deictic etymology (see Ehlich 1994 for the concept of a "functional etymology"). If the term is to be applied cross-linguistically to also include verbal suffixes used to derive verbal nouns and participles, this implies a conceptual expansion of the term.

---

the reasons to analyse *diye-* and *dep*-constructions as quotative forms, which would exclude them from a discussion of complementizers (Jochen Rehbein, p.c., 21 September, 2013). We exclude these structures from the present discussion (but see Herkenrath & Karakoç 2007a [2002]) in order to focus on the functions of the bound morphemes in non-finite complement clauses.

4 In citing data from the spoken language corpora ENDFAS and SKOBI (see Section 3 for more detail), we give the file name to document the precise place of the finding, as the corpora do not make use of any page numbering.

5 See Rizzi (1997) for a theoretical model that distinguishes the CP from the IP area mainly on the criterium of morphological affixation versus morphologically independent words.

The second level of our approach is comparative, confronting two Turkic languages, one Oghuz (Turkish) and one Kipchak (Noghay). Can the two languages be shown to have complementation markers that semantically function in more or less identical ways? Or, are there any interesting language-specific differences? If so, how can these differences be theoretically modelled in terms of the functional micro-structure of the individual forms?

Thirdly, we will be comparing the functional semantics of individual forms within the system of one given language, e.g. Turkish *-DIK* versus Turkish *-mA* or Noghay *-GAn* versus Noghay *-(U)w*.[6] At this level, we expect to find interesting differences between forms deriving from aspectual markers versus from derivational affixes. It is at this level that we also try to look at empirical or, where lacking, constructed minimal pairs in order to tease out language-internal semantic oppositions between individual forms.

The paper is organized as follows: After a brief portrait of the two languages under discussion (Section 2), and a brief description of our data (Section 3), we discuss some problems related to assigning complementizer status in these two languages (Section 4). This discussion is presented in two main strands, first, in relation to a general typological discussion, and second, in a way more internally related to Turkic linguistics. Section 5 presents some morphosyntactic aspects of clausal complementation in Turkish and Noghay, focusing on non-finite constructions and presenting the different types of subordinators as well as the inner and outer syntactic relations of the assertive and interrogative constructions embedded by them. Section 6 attempts to tease out some semantic and functional differences between the individual morphemes in question. Section 7 shows some other subordinating and non-subordinating functions of the same morphemes.[7]

2 Languages investigated

Turkish and Noghay are two genealogically related languages. As members of the Turkic language family, they share common structural features. For instance, the grammatical categories are mainly expressed by bound morphemes exhibiting alternating allomorphs in complementary distribution.

6 We cannot make any serious attempt at discussing the discourse characteristics of complement clauses or of matrix-complement constructions in this paper. See e.g., Rehbein (2007) for a functional-pragmatic analysis and Boye & Harder (2007) for a specific cognitive approach.

7 The present paper focuses on the non-finite suffixes, for the most part leaving aside a discussion of the functional (or semantic) role of the morphemes they are combined with (such as possessives, case markers, and postpositions).

Turkish, the official language of Turkey, belongs to the Oghuz or southwestern branch of the Turkic languages. This branch also includes Azerbaijanian, spoken in Azerbaijan and Iran; Gagauz, spoken particularly in Moldova; Turkmen, spoken in Turkmenistan and Iran; Khorasan Turkic, spoken in Iran; as well as different Turkic dialects spoken in southern Iran and Afghanistan. Turkish is spoken by about 75 million people in Turkey (according to the 2012 census). Outside Turkey, there are also Turkish speakers, particularly in the Balkans, Cyprus, and Western European countries.

Noghay is a member of the Kipchak or northwestern branch of Turkic languages. Within this branch it forms, together with Kazakh, Kirghiz, Karakalpak and Kipchak Uzbek, a southern subgroup (South Kipchak) that is also called the Aralo-Caspian group (see Csató & Karakoç 2006 [1998]). Nowadays, Noghays are numerically a small group of Turkic people who mainly live in the northern Caucasus dispersed over different autonomous republics and administrative districts of Russia. The number of Noghays in the northern Caucasus amounts to about 103,000 (according to the 2010 census) (Karakoç 2013a).

3 Data

Throughout this paper, the status of examples is to illustrate different kinds of constructions without following any quantitative goals. The Turkish data are both spoken and written data, the spoken data having been taken from the ENDFAS/SKOBI corpus.[8] The data of written Turkish comes from contemporary novels. Occasional examples were randomly taken from non-fiction textual sources; this method was sometimes used in order to find examples illustrating a specific functional contrast. The Noghay examples are mostly from literary sources; the examples of spoken Noghay are taken from Karakoç (2005).

In what follows, Turkish examples will be presented in the official orthography of Turkey. Noghay is written in the Cyrillic alphabet and the data will be rendered in Turcological transcription (see Karakoç 2005). Examples are given in

8 ENDFAS (The development of narrative discourse abilities in German and Turkish in family and school) and SKOBI (Linguistic connectivity in Turkish-German bilingual children) are two research projects that were sponsored by the German Science Foundation and supervised by Jochen Rehbein at the University of Hamburg (Institut für Germanistik I and Cooperative Research Centre 538 Multilingualism). See Rehbein (2009), Rehbein et al. (2009), Herkenrath & Rehbein (2012). These corpora contain two sets of spoken language recorded in monolingual and bilingual family constellations and transcribed according to the HIAT conventions (Ehlich & Rehbein 1976; Rehbein et al. 2004). We will generally refer to the monolingual control data from Turkey in the present paper.

italics. Complement clauses in the examples are highlighted in square brackets. In glossing the complementizing morphemes, we only name the item, e.g., DIK, YACAK, MA, MAK, YIŞ, leaving any categorization to the analysis.

A dash after a word indicates a verbal stem, e.g., *gel-* 'come'. A dash in front of a morpheme denotes a suffix, e.g., *-DIK* (a participle suffix). Asterisks (*) denote unacceptable constructions. Capital letters show morphophonemes:

A Alternation of non-high vowels *a* and *e* in Turkish
Alternation of non-high vowels *a*, *e* in standard written Noghay. *A* is realized as a palatal glide *y* after stem-final vowels in Noghay

I Fourfold alternation of high-vowels *ı*, *i*, *u* and *ü* in Turkish
Twofold alternation of high-vowels *ï* and *i* in standard written Noghay

D Alternation of *d* and *t* in Turkish
Alternation of *d*, *t* and *n* in Noghay

K Alternation of *k* (phonologically *k* or *ḳ*) and *ğ* (phonologically *Ø*) in Turkish

G Alternation of *g*, *γ*, *k* and *ḳ* in Noghay

M Alternation of *m*, *b* and *p* in Noghay

Segments in parentheses show the phonemes that are only realized in certain environments.

Example (2) illustrates how the subordinating morpheme *-DIK* in Turkish undergoes a morphophonemic variation depending on the vocalic and consonantic givens of the previous and subsequent syllables. The example displays six of the sixteen possible forms of *-DIK*: *-tık, -tik, -tuk, -tük, -dık, -dik, -duk, dük, -tığ, -tiğ, -tuğ, -tüğ, -dığ, -diğ, -duğ, -düğ*:

Turkish

(2) *Kadın,* [*koca-sın-ın* *dün* *ne-ler* *ye-diğ-in-i*],
woman husband-POSS.3-GEN yesterday what-PL eat-DIK-POSS.3-ACC
[*ne-ler* *konuş-tuğ-un-u*], *dükkan-dan* *ev-e* *gel-ir-ken*
what-PL speak-DIK-POSS.3-ACC shop-ABL house-DAT come-AOR-CVB.COB
kuş-lar-a *gökyüzün-de-ki* *nasıl* *bak-tığ-ın-ı,* [*küçük*
bird-PL-DAT sky-LOC-ADJ how look-DIK-POSS.3-ACC little
kız-ın-ı *eşik-te* *neden* *tokatla-dığ-ın-ı*] *ve* [*ruh-um*
girl-POSS.3-ACC threshold-LOC why slap-DIK-POSS.3-ACC and soul-POSS.1SG
daral-ıyor *diye* *içini çek-tiğ-in-de* *sırt-ın-da* *hangi* *gömleğ-in*
get_tight-PRS saying sigh-DIK-POSS.3-LOC back-POSS.3-LOC which shirt-GEN
bulun-duğ-un-u] *anlat-mış-tı* *sonra*.
be_found-DIK-POSS.3-ACC tell-POST-PST.COP later
'Then the woman had told what her husband had eaten yesterday, what he had talked about, how he had looked at the birds in the sky when coming home from the shop, why he had slapped his little daughter on the threshold and what shirt he had been wearing when he sighed, signifying that he was fed up.'
(Toptaş 1995: 18)

4 'Complementizer' status of the subordinating suffixes

This volume employs the term "complementizer" to refer to linguistic elements or expressions that conform to Noonan's (2007 [1985]: 55) definition: "a word, particle, clitic or affix, one of whose functions it is to identify [a clause] as a complement". This concept refers to a subcase of the more general concept of "complementation markers", a concept that includes entire constructions, word orders, prosodic features, etc. The present volume uses the three terms, "canonical complementizers" (the most specialized concept, where the predicate of the subordinate construction is indeed finite and the subordinator is a word rather than a suffix), "complementizers" and "complementation markers" (the most general concept), as subcases of each other (see also Kehayov & Boye, this volume). Although we employ the cross-linguistic terminology, we also want to keep an eye on the question of how a concept such as "complementizer" can be applied to Turkic without changing the connotations with which it has been used in the literature. It may be the case that employing the term "complementizer" too freely would obscure distinctions that might turn out to be not just typological but more deeply functional.

This section is thus intended as a brief theoretical preliminary to the language-specific discussions in Section 5 onwards. It aims to relate the typological literature that serves as a theoretical basis for this volume to the specific situation in Turkish and Noghay.[9] By way of this theoretical discussion, we hope to observe the caveat that what is intuitively a complementizer need not be what actually identifies the complement as a complement.

There exists an interesting typological discussion, offering a variety of definitions of the concept "complementizer", some of which seem broad enough to include morphological subordinators of the kind found in Turkish and Turkic, and others, narrower, that would rather seem to exclude such devices and rather analyse them as "complementation strategies". The discussion also involves the issue of whether some clause types are nominals instead, and of how to draw a line between clause-type and nominal-type constructions.

The concept of "complementizer" was first introduced by Rosenbaum (1967) and the discussion continued with Lees (1968 [1963], 1965), Chomsky (2005

9 The problems we are confronted with basically concern two questions: first, whether or not it is justified to speak of "complementizers" in Turkish and Noghay in the first place, and second, if we do ascribe complementizer status to them, whether such status should be ascribed to particular suffixes on their own or rather to combinations of them.

[1980], 1993 [1981], 1986, 1995), Abney (1987), Szabolcsi (1992), Bernstein (2001), and others. Noonan's (2007 [1985]: 55) concept, as already mentioned, also offers a broad definition of a complementizer: "a word, particle, clitic or affix, one of whose functions it is to identify [a clause] as a complement". This definition can include, for our purposes, morphologically bound subordinators of the kind we find in Turkish and Noghay. Noonan includes in his inventory complementizers from a range of languages, deriving from linguistic elements as diverse as demonstrative, interrogative and relative pronouns, conjunctions, adpositions, case markers (affixed to finite verb forms or preposed as morphologically unbound particles), or verbs (comparable to Turkish *diye* [say-CVB]). These elements can be morphologically free or bound, and can occur in a variety of syntactic positions (Noonan 2007 [1985]: 57–59). Noonan's concept includes non-finite as well as finite constructions, varying degrees of sentence-likeness, and reductions in categorical distinctions (mostly in tense and sometimes aspect, but not in actant marking). Subjects may be treated differently in complement clauses (e.g., with respect to the assignment of subject case) or may be unspecified (as in infinitive constructions). None of these differences have any consequences for the complementizer status of the elements that connect the complement to the matrix clause (Noonan 2007 [1985]: 60–61).

While these typological models apply a broad understanding of complementizers, others, such as those proposed in Dixon (1995, 2009), given the way they analyse constructions from a variety of typologically diverse languages, would rather exclude the Turkic devices. Dixon (1995, 2006) draws a fundamental distinction between "complement clauses" on the one hand and "complementation strategies" on the other. Several of the articles (Deutscher 2006; Aikhenvald 2006; Hellwig 2006; Fleck 2006) in Dixon & Aikhenvald (2006) discuss nominalizations as complementation strategies, i.e., as linguistic means not fully equivalent to complementizers. Among all these different cases, the constructions in Matses (Fleck 2006) – with nominalizing strategies for clausal constructions in argument positions, presented as complementation strategies, i.e., not as complement clauses – seem to come closest to the structures we are going to describe for Turkish and Noghay. Cristofaro's (2005) functional-typological approach speaks of complementizers in referring to elements that occur at the beginning of a clause, introducing it. They derive from nouns and pronouns, i.e. deictic expressions or relative pronouns (with a resumptive function), a situation more or less typical of Indo-European.

A recent typological account of the phenomena under discussion is given by the contributions in Kornfilt & Whitman (2012a), which make very scarce use of the term "complementizer". Where they do so, it is usually with respect to morphologically unbound expressions such as *that* in English or *dien* in the Turkic

language Sakha (a cognate of Turkish *diye*, Baker 2012). With respect to the procedures under discussion in the present article, they speak of "nominalization". Kornfilt & Whitman (2012c) systematize this into a "Functional Nominalization Thesis"; this is clearly an alternative to expanding the concept of complementizer.

The morphosyntactic status of Turkish complement clauses as lying somewhere in between nominal, clausal and verbal has led to a wide debate, particularly in generative syntax theory, starting with Borsley & Kornfilt's (2000) model of "mixed extended projections" and leading up to Kornfilt & Whitman's (2012c) concept of a "terminus of the verbal projection in nominalizations". Basically, the more arguments that can be found for a verbal status of the constructions in question, the more it seems to make sense to speak of complementizers instead of simply nominalizers. In what follows, we briefly summarize the main arguments for both sides.

The most striking nominal characteristic of subordinated predicates is their combinability with possessive and case suffixes, i.e. nominal suffixes (Kornfilt 1997: 46, 2006, 2007 and elsewhere): "[...] in Turkish, most complements are nominalized; these have no subordinating conjunctions in the Indo-European sense. One might view the nominalization suffixes as markers of subordination" (Kornfilt 1997: 15, 45). The nominalization markers occupy the position of tense, according to Kornfilt (1997: 46; see also Kornfilt & Whitman 2012c: 1301); this has some semantic implications, as we shall see later on. Other arguments for a nominal status are the syntactic position occupied by complement clauses, which is the same as that of ordinary nominal complements (Kornfilt 1997: 46). Kennelly (1993, 1994) specifically discusses *-DIK*-constructions in light of the DP hypothesis of generative syntax theory.

The work of Johanson (e.g., 1975a, 1990a, 1990b, 1996, 2010, 2013) shows in detail what distinguishes Turkic nominalized propositions from Indo-European abstract nominalisations. Stressing the verbal character of non-finite complement clauses, Johanson (1975a: 110) sets them off against abstract nominalizations of the kind also found in English (e.g., *the man's arrival*), also showing how Indo-European nominalization occurs at the expense of tense marking, *genera verbi*, and argument structure to a greater degree than is the case in Turkish nominalized propositions. Functionally and sentence-hierarchically, Johanson (1973, 1975a, 1975b, 1990a) identifies the said non-finite constructions in Turkish with grammatically dependent subordinate clauses in those Indo-European languages.

Theoretical models of "hybrid" constructions or "mixed extended projections", based on generative syntactic categories, can be seen as ways of reconciling the nominal and the verbal side: see Baker's (1986, 1988) concept of "incor-

poration" or Borsley & Kornfilt's (2000) model of a "mixed extended projection". Other hybrid accounts are proposed by Kornfilt (2007), looking at the question of finiteness in an abstract manner that includes different forms of agreement, as well as Kornfilt & Whitman (2012b, 2012c), whose "Functional Nominalization Thesis" integrates different types of nominalization occurring at different syntactic levels. Turkish, in this view, instantiates TP nominalizations, to the extent that it is at the level of tense that the verbal projections give way to nominal ones, possibly with a difference between *-mA* and *-DIK* (Kornfilt & Whitman 2012c: 1302).

In just very briefly sketching these debates, we hope to have presented a basic idea of how some linguists might accept complementizer status for the subordinating morphemes, whereas others deny this possibility and speak of nominalization or non-finitization (with subtle differences between the two) instead. As we now turn to our data-based descriptions and analyses, we hope to proceed with great caution.

5 Some morphosyntactic properties

5.1 Inventory of complementation markers

In Turkish and Noghay, non-finite complement clauses are based on verbal-nominal or participial complementation markers; there are also complex forms that express aspectual distinctions.

1. The Turkish verbal-nominal markers are *-mAK*, *-mA* and *-(y)Iş*, and those in Noghay are *-(U)w* and *-MAGA*[10] (see Table 1). The Turkish form *-mAK* and the Noghay form *-MAGA* do not take possessive suffixes and only occur in control constructions. The Turkish form *-mA* and the Noghay form *-(U)w* can occur both without possessive markers in control constructions and with possessive markers in constructions having their own subjects. The phenomenon of obligatory control and thus the functions of Turkish *-mAK* and *-mA*, as well as Noghay *-MAGA* and *-(U)w*, in control constructions will not be further discussed in this paper. See Karakoç (2013b) for an analysis of obligatory control in Turkish and Noghay complement clauses.

10 Diachronically viewed, the marker *-MAGA* in Noghay has developed from a combination of the infinitive in *-MA* plus the dative case *-GA*. The infinitive often contains a dative marker in other Turkic languages as well. Haspelmath (1989) shows that the infinitives and similar complement forms of various languages have historically developed from the purposive forms containing verbal nouns in dative, locative or other cases (see also Karakoç 2013b).

Table 1: Verbal-nominal markers in Turkish and Noghay

Verbal-nominal markers				
Turkish			**Noghay**[11]	
-mAK without possessive; only in control constructions	*-mA* with or without possessive	*-(y)Iş* with or without possessive	*-MAGA* without possessive; only in control constructions	*-(U)w* with or without possessive

2. Participial morphemes have either a simple or a complex (postverbial) morphological structure: simple markers cannot be morphemically segmented. These are *-DIK* and *-(y)AcAK* in Turkish and *-GAn* and *-(A)yAK* in Noghay. Postverbial markers are only found in Noghay: *-Atayan*, *-(A)yatḳan*, *-(A)yatïryan* (Table 2). Diachronically, these forms are developed from postverbial constructions, see Karakoç (2005, 2007a [2002], 2007b).

Table 2: Participial markers in Turkish and Noghay

	Participial markers	
	Turkish	**Noghay**
Simple forms	*-DIK*	*-GAn*[2]
	-(y)AcAK	*-(A)yAK*
Postverbial forms		*-Atayan*
		-(A)yatḳan
		-(A)yatïryan

3. Some complementation markers consist of finite markers and copular forms. In Turkish, the finite aspectual forms *-(Ø)Iyor, -mAktA* and *-mIş* are combined with the copular verb *ol-* 'become, be'. The verbal-nominal or participial forms mentioned above attach to this copular verb: *ol-duK-*, *ol-ma-*, *ol-mak,*

11 The Noghay suffixes *-(I)s* and *-(I)š* as well as *-MAK* and *-MA*, which are cognates of Turkish *-(y)Iş, -mAK* and *-mA* respectively, are exclusively used as derivational suffixes.

12 The Noghay participle in *-GAn* can sometimes occur in non-possessive complement clauses denoting a subject control relation (see Karakoç 2007a [2002]: 353, 2013b: 319).

ol-uş-, ol-acaK-. In Noghay, the aspectual markers *-AdI, -(A)r* and *-GAn* are combined with the copular marker *eken* (Table 3):[13]

Table 3: Complex markers in Turkish and Noghay

Complex complementation markers including copular forms			
Turkish		**Noghay**	
-(Ø)Iyor *-mAktA* *-mIş*	*+ ol-* (*+-DIK/ -(y)AcAK/ -mA/ -(y)Iş*)	*-(A)r* *-AdI* *-GAn*	*+ eken-*

5.2 Inner structure of complement clauses

The inner structure of complement clauses relates to argument structure, subject marking, and types of predicates. It has the same word order as in finite clauses. Turkish and Noghay complement clauses have been analysed as containing "deranked" verb forms that can only be used in subordinate clauses and "lack some categorial distinctions relevant to main clauses" (Johanson 2013: 75; see also Csató 2010). The subject markers in finite verbal predicates and non-finite complement clauses are of a different nature. For an illustration of the unfolding of an argument structure inside nominalized complement clauses, we refer the readers to example (2) in Section 3. The present section looks at subject marking and types of complement-taking predicates. In Turkish, a complement clause with a full subject noun phrase occurs in the form of a genitive-possessive construction. As in (3), the overt subject, which carries a genitive marker, i.e. *ben-im* [DEI.1SG-GEN] 'my', fulfils the role of possessor and the predicate exhibiting possessive morphology, i.e. *gel-diğ-im-i* [come-DIK-POSS.1SG-ACC] 'that I come/ have come/ came', assumes the role of a possessed element, as in (3):

13 There are also some other nominalized forms, less frequently used and often simply noun-like, that may still unfold an argument structure, such as Turkish *-mIşlIK* and *-mAzlIK* as well as Noghay *-GAnlIK*. For reasons of space, and since these forms require a more detailed analysis, we do not include them in this paper.

Turkish

(3) [“*Hadi* *di-yelim* *çırağ-ı* *fark_et-me-di-n,* *ama*
come_on say-OPT.1PL apprentice-ACC notice-NEG-PST-2SG but
[*sen-den* *sonra* *bura-ya* *benim* *gel-diğ-im-i*] *de*
DEI.2SG-ABL after here-DAT DEI.1SG.GEN come-DIK-POSS.1SG-ACC also
mi *gör-me-di-n,”* *diye* *sor-ma-m-a*] *gerek* *yok-tu.*
Q see-NEG-PST-2SG saying ask-MA-POSS.1SG-DAT need nonexistent-PST.COP
‘There was no need for me to ask (or: that I ask): “Come on, let’s say you didn’t notice the apprentice, but didn’t you see that I came here after you, either?”’
(Toptaş 1995: 111)

Subject noun phrases can also appear in the nominative; these occurrences seem to be much rarer in Turkish and mainly concern subjects with a non-specific reference (see Erguvanlı Taylan 1994; Göksel & Kerslake 2005: 430–431; Johanson 2006, 2013). (4) illustrates a nominative subject with a non-specific reference:

Turkish

(4) [*Masa-m-ın* *çekmece-ler-in-de* *gizlilik* *derece-si*
table-POSS.1SG-GEN drawer-PL-POSS.3-LOC secrecy degree-POSS.3
HİZMETE *ÖZEL-den* *yüksek* *evrak*
official_use-DAT reserved-ABL high document
bulun-ma-dığ-ın-ı] *taahhüt* *ed-er-im.*
be_found-NEG-DIK-POSS.3-ACC commitment make-AOR-1SG
‘I undertake a commitment that there is no document with a degree of secrecy higher than FOR OFFICIAL USE ONLY in the drawers of my desk.
(Kaygusuz 2006: 10)

The subject in Noghay complement clauses mostly appears in nominative case, even in contexts in which, translated into Turkish, they would have to be genitive-marked (5). See also example (17) below.

Noghay

(5) *Süyün-ö-müz* [*biz-im* *awul-umuz-ya* *ḳonak-lar* *kel-gen-in-e*].
be_happy-PRS-1PL DEI.1PL-GEN village-POSS.1PL-DAT guest-PL come-GAN-POSS.3-DAT
‘We are happy that the guests have come to our village.’
(Karakoç 2005: 216)

The predicate of a complement clause can be 1) a lexical verb, 2) a noun + a copular verb, 3) an adjective + a copular verb. If the predicate is not verbal, a copular verb has to be added to host the subordinating morphology. Example (6) contains a predicate consisting of an adjective and the copular verb *ol-* (for the various semantic readings of the copular verb *ol-* in Turkish complement clauses, see Karakoç 2007a [2002]):

Turkish

(6) *Bi de bi yer var, ora-nın su-lar-ı • sıcak ee*
one also one place existent there-GEN water-PL-POSS.3 hot IJ
[*neden sıcak ol-duğ-un-u*] *bil-emi-yor-um.*
why hot be-DIK-POSS.3-ACC know-MOD.NEG-PRS-1SG
'And there's a place, its waters are • hot eh, why they are hot, I don't know.'
(Rehbein 2009: EFE05tk_Can_m_0395_5_ENF)

In Noghay, if the predicate is a noun or an adjective, the static copular particles *eken* or *ekenlik* are employed (7) (for the other copular forms used in Noghay complement clauses, see Karakoç 2007a [2002]):

Noghay

(7) *Ḳïsḳayaḳlï Farida-dï tanï-y-dï [bay-ï-nïŋ ana-sï*
woman Farida-ACC recognize-PRS-3SG husband-POSS.3-GEN mother-POSS.3
eken-i-n] de aŋla-y-dï.
EKEN-POSS.3-ACC also understand-PRS-3SG
'The woman recognizes Farida and also understands that she is her husband's mother.'
(Karakoç 2005: 223)

5.3 Integration of complement clauses into a superordinate structure

Within the superordinate clauses, depending on the argument structure of the predicate, complement clauses assume various syntactic roles, such as: 1) subjects (nominative), 2) direct objects (accusative), 3) indirect objects (dative), 4) oblique objects (locative, ablative, instrumental), 5) postpositional objects, and 6) predicate nouns in copular clauses. Example (8) illustrates the use of a complement clause in the role of subject, (9) in the role of direct object, and (10) in the role of oblique (ablative) object:

Turkish

(8) *Aá Ne güzel. [Balığ-ı sev-iyor ol-ma-n] işte hoş birşey.*
IJ what nice fish-ACC like-PRS be-MA-POSS.2SG PTC nice one thing
'Oh, how nice. That you like fish, you know, is really nice.'
(Rehbein et al. 2009 : EFE07tk_Muh_b_0821_5_SKO)

Turkish

(9) *Hemen ertesi gün köy-den geç-en bir çerçi,* [*Nuri-nin mavi bir*
at_once next day village-ABL pass-PTCP one peddler Nuri-GEN blue one
kamyon-da şoförlük yap-tığ-ın-ı] *söyle-miş-ti.* [...] *Aslında*
lorry-LOC driving make-DIK-POSS.3-ACC say-POST-PST.COP in_fact
[*çerçi-nin köy-e ne zaman gel-ip git-tiğ-in-i*]
peddler-GEN village-DAT what time come-CVB go-DIK-POSS.3-ACC
bil-en yok-tu.
know-PTCP nonexistent-PST.COP
'The very next day, a peddler passing through the village said that Nuri was (or: had been) a driver in a blue lorry. [...] In fact, there wasn't anybody who knew when and how often the peddler came/had come to the village.'
(Toptaş 1995: 32)

Turkish

(10) *Hızla yüksel-ir-ken,* [*sırr-ı-nın açığa_çık-ma-sın-dan*] *endişe*
speed-INS rise-AOR-CVB.COB secret-POSS.3-GEN come_out-MA-POSS.3-ABL anxiety
duy-ma-mış-tı.
feel-NEG-POST-PST.COP
While he was moving up quickly, he had not felt any anxiety about his secret coming to light.'
(Şafak 1999: 23)

5.4 Interrogative complement clauses

In Turkish, *wh*-clauses are embedded by the same morphological means as non-interrogative clauses, and the *wh*-element remains unaffected by the subordinating procedures. The embedding of propositional questions differs more substantially from the corresponding non-embedded propositional questions. While the latter contain a cliticized Q-particle, which can occur in different syntactic positions as a function of expressing interrogative focus, the embedded constructions express interrogativity by means of two adjacent forms of the same verb: a converbial positive and a participial negative form. The following passage of discourse from a conversation on how to eat *çiğ köfte* 'raw meatballs' (11) contains both *wh*- and polar interrogatives in embedded and non-embedded constructions:

Turkish

(11) a. Ken: *Nasıl yap-ıyo-lar, kim yap-ıyo siz-de çiğ_köfte-yi?*
how make-PRS-3PL who make-PRS DEI.2PL-LOC çiğ_köfte-ACC
'How do they prepare it, who makes it in your family, the çiğ köfte?'

b. Sed: *Baba-m.*
father-POSS.1SG
'My Dad.'

c. Ken: *Nasıl yap-ıyo baba-n?*
how make-PRS father-POSS.2SG
'How does he do it, your father?'

d. Sed: *Çiğ_köfte-yi kız-lar yap-amaz ki!*
çiğ_köfte-ACC girl-PL make-MOD.NEG.AOR PRT
'Girls don't make *çiğ köfte*, man!'

e. Ken: *Öyle mi, erkek-ler mi yap-ar on-u?*
so Q boy-PL Q make-AOR DEI-ACC
'That so, is it boys who make it?'

f. Sed: *Hm̆˙*
IJ

g. Ken: *Nasıl yap-ıl-ıyo bi düşün?*
how make-PASS-PRS once think
'How is it made, think about it?

h. *Nasıl yap-ıl-dığ-ın-ı bi anlat bak-alım hadi!*
how make-PASS-DIK-POSS.3-ACC once tell look-OPT.1PL come_on
'Just tell me once how it is made!'

i. Sed: *Ya ben de bil-m-iyor-um ki, hiç izle-me-di-m ben*
PTC I also know-NEG-PRS-1SG PTC at_all watch-NEG-PST-1SG I
• *nası yap-ıl-dığ-ın-ı!*
how make-PASS-DIK-POSS.3-ACC
'I don't know either, I never watched • how it is made!'

j. Ken: *((1s)) Sadece ye-di-n!*
only eat-PST-2SG
'((1s)) You've only eaten it!'

k. Sed: *Hmm̀˙*
IJ

l. Ken: *Nasıl ye-n-iyo bi anlat!*
how eat-PASS-PRS once tell
'How is it eaten, tell me!'

m. Sed: *Ya işte marul-un iç-in-e koy-uyo-sun ya da*
PTC PTC lettuce-GEN inside-POSS.3-DAT put-PRS-2SG or
böyle bi marul gibi bişey-in iç-in-e
like one lettuce like one_thing-GEN inside-POSS.3-DAT
koy-uyo-sun, yi-yo-sun!
- put-PRS-2SG eat-PRS-2SG
'Well okay, you put it into a piece of lettuce, or you put it like into a lettuce-like thing, you eat it!'

n. Ken: *O kadar!*
that degree
'That's it!'

o. Sed: *Hmhm̀˙ El-in-le.*
IJ hand-POSS.2SG-INS
'IJ With your hands.'

p. Ken: *İç-in-e* *limon,* *tuz* *muz falan* •• *at-m-ıyo*
inside-POSS.3-DAT lemon salt and_so_on throw-NEG-PRS
mu-sun?
Q-2SG
'Don't you • • add lemon, salt or something, and so on?'

q. Sed: *Hep/* • • *biraz* *limon* *at-abil-ir-sin!*
all a_little lemon throw-MOD-AOR-2SG
'All/ • • you can add a little lemon!'

r. Ken: *Biraz* *limon* *sık-abil-iyo-sun.*
a_little lemon squeeze-MOD-PRS-2SG
'You can squeeze a little lemon.'

s. • *Háhà* *on-dan* *sonra.*
IJ DEI-ABL after
'• I see, after that.'

t. Sed: *On-dan* *sonra* *da* *mide-n-e* *in-iyo.*
DEI-ABL after also stomach-POSS.2SG-DAT go_down-PRS
'And after that it goes down to your stomach.'

u. Ken: *Peki* *güzel* *bişey* *mi?*
okay nice one_thing Q
'Okay, is it nice?'

v. Sed: *Çok* *güzel!*
very nice
'Very nice!'

w. Ken *Acı* *diğil* *mi?*
hot not Q
'Isn't it hot?'

x. Sed: *((2s))* *Ooó* *şey* *malzeme-sin-e* *bağlı,* *acı* *ol-up*
IJ thing ingredient-POSS.3-DAT dependent hot be-CVB
ol-ma-dığ-ı.
be-NEG-DIK-POSS.3
'((2s)) Ooh, hang on, it depends on the ingredients whether or not it's hot.'

(Rehbein et al. 2009: EFE02tk_Sed_m_0747_SKO)

As can be seen, non-embedded *wh*-constructions contain finite predicates, such as in segments a, c and g: *Kim yapıyor sizde çiğ köfteyi?* 'Who prepares *çiğ köfte* in your family?', *Nasıl yapıyo baban?* 'How does he do it, your dad?' and *Nasıl yapılıyor bi düşün?* 'How is it made, think about it'. The corresponding embedded constructions, where the *wh*-question becomes the complement of a *verbum dicendi*, are non-finite: *Nasıl yapıldığını bi anlat bakalım hadi!* 'Just tell me once how it is made, come on!' and: *Ya ben de bilmiyorum ki, hiç izlemedim ben • nasıl yapıldığını!* 'Hey, I don't know either, I never watched how it is made!' (in segments h, i).

Propositional questions can be studied a little further down in the same passage: *İçine limon, tuz muz falan* • • *atmıyor musun?* 'Don't you • • add lemon, salt or anything, and so on?' (segment p). Still further down, there are two occur-

rences of uninflected *mI* in unembedded nominal sentences: *Peki güzel bişey mi?* 'Okay, is it nice?' and *Acı diğil mi?* 'Isn't it hot?' (segments u, w). The corresponding subject clause in segment x, *Ooó şey malzemesine bağlı, acı olup olmadığı* 'That, well, depends on the ingredients, whether it is hot' is a non-finite construction made up of the converbial *olup* 'be-CVB' and the negated participial *olmadığı*.

6 The functional quality of the bound morphemes

This section attempts to describe the functional quality of the complementizing morphemes in both languages, and finds some contrasts at the intralingual and interlingual levels. With reference to ongoing Turcological discussions, our analysis presents some functional-semantic concepts expressed by the individual forms, focusing on the expression of knowledge-related and aspecto-temporal ideas. This also implies looking at the interaction of the subordinating morphemes with the lexical semantics of matrix predicates. Methodologically, we adopt a data-driven approach, looking at semantic oppositions and issues of interchangeability and applying some minimal-pair testing to empirically obtained examples. On the basis of the semantic phenomena teased out in this way, we draw some links to the theoretical discussion above.

6.1 *-DIK* versus *-mA* in Turkish and *-GAn* versus *-(U)w* in Noghay

The following Turkish examples (12) and (13) differ with respect to the bound morphemes, *-DIK* and *-mA*, used in the embedded clauses, but share the same matrix verb, *sevin-* 'be happy, glad'. While matrix verbs such as *sevin-*, which are semantically compatible with all four complementation markers (*-DIK*, *-mA*, *-(y)Iş*, and *-(y)AcAK*), are relatively rare, they are interesting in that they allow an analysis of the semantic microstructure of the morphemes. See also the examples (28) and (34), in which the matrix verb *sevin-* contains a *-(y)Iş* and a *-(y)AcAK* based complement clause respectively.-

Turkish

(12) *O-nun* *gibi* *bi* *kurdele-si* *var-mış.* *Okul-dan*
DEI-GEN like one ribbon-POSS.3 existent-EVID.COP school-ABL
gel-iyo-muş. *Anne-si* [*bun-u* *al-dığ-ın-a*]
come-PRS-EVID.COP mother-POSS.3 DEI-ACC receive-DIK-POSS.3-DAT
çok *sevin-miş.*
much be_happy-POST
'She had a ribbon like that. She came from school. Her mother was very happy about the fact that she received this.'
(Rehbein et al. 2009: EFE02tk_Hai_m_0790_3_SKO)

Turkish

(13) [*Okul-un* *tatil* *ol-ma-sın-a*] *sevin-iyo-sun* *yaa.*
school-GEN holiday be-MA-POSS.3-DAT be_happy-PRS-2SG PTC
'Yeah, you're happy about your school being on holiday.'
(Rehbein et al. 2009: EFE08tk_Sed_m_0602_3_SKO)

Linguists have time and again tried to tease out the functional opposition between the complementation markers *-DIK* and *-mA*. The idea that the distinction between the participial and verbal-nominal complementation markers is characterized by means of the concepts factive versus non-factive goes back to Lees ([1963]1968, 1965). In Lees' view, complement clauses based on *-DIK* are characterized as "factive nominalizations", while the clauses based on *-mA* are analysed as correlating to an action meaning ("action noun" or "light infinitive nominalization"). This conceptual distinction can also be found in subsequent descriptions and grammars of Turkish, e.g., those of Underhill (1976) and Kornfilt (1997). Van Schaaik (1999), working within a Functional Grammar framework, makes a distinction between "act" (i.e. *-mA*, forming a predicational clause, expressing a state-of-affairs) and "fact" (embedding a proposition, i.e. *-DIK*), thereby linking this distinction to the semantics of specific matrix verbs. In this view, matrix verbs are classified into three groups: verbs that obligatorily combine with *-DIK*, others that obligatorily combine with *-mA*, and a third group, in which the choice between subordinating suffixes is somewhat free, allowing the expression of semantic differentiations.

According to Csató (1999) complement clauses based on the participle in *-DIK* have an "assertive" (indicative) meaning, whereas the *-mA* clauses are "non-assertive": *-DIK* clauses have illocutionary force and convey propositions with a possible truth value. *-mA* clauses express predications without any possible truth value. Csató shows how predications involving *-DIK* or *-mA* are typically employed together with semantically defined classes of matrix predicates. Assertive predicates are usually constructed with *-DIK*. *-mA*, while it also can be com-

bined with assertive predicates, requires another interpretation. In this regard, *-mA* describes the action and appears in the unmarked, neutral form, combinable with most matrix predicates. Whereas matrix predicates expressing a speaker's stance towards the validity of a proposition can occur with *-DIK* complements, predications with *-mA* are neutral with respect to whether a verbalized event has actually occurred or not (cf. van Schaaik 2014).

Applying a semantic typology proposed by Noonan (2007 [1985]), Csató (2010: 66–67) defines semantic groups of predicates, suggesting selectional tendencies and frequency patterns. The first group of predicates, which mainly head complement clauses based on *-DIK*, contains: 1) utterance predicates, e.g., *söyle-* 'say', *sor-* 'ask', *anlat-* 'tell', *bildir-* 'make known', 2) predicates expressing "propositional attitudes specifying the attitude of a person in relation to the possible fact designated by the propositional complement" (Csató 2010: 66), e.g., *emin ol-* 'be sure', *inan-* 'believe that it is true', *doğru ol-* 'be true', *yanlış ol-* 'be false', 3), predicates of knowledge, acquisition and loss of knowledge, immediate perception, e.g., *anla-* 'understand', *bil-* 'know', *göster-* 'show', *gör-* 'see', *duy-* 'hear', *farket-* 'notice', *hatırla-* 'remember', 4) predications of mental perception, e.g., *pişman ol-* 'regret, be sorry about' (Csató 2010: 66). The second group of predicates, which govern clauses based on the verbal noun *-mA*, includes: 1) predicates of practical manipulation, e.g., *emret-* 'order, command', *zorla-* 'oblige', *engelle-* 'prevent', *izin ver-* 'allow', *öğütle-* 'advise', 2) volitional predicates, e.g., *iste-* 'want', *rica et-* 'request', *arzu et-* 'desire', *um-* 'hope', 3) achievement predicates, e.g., *başar-* 'accomplish, achieve', *çalış-* 'try', 4) phasal predicates, e.g., *başla-* 'start, begin', *bitir-* 'complete, 5) commentative predicates, e.g., *memnun ol-* 'be pleased', *sevin-* 'be glad', *üzül-* 'be sorry', *şaşır-* 'be surprised', *kork-* 'be afraid, fear', 6) objective modal predicates, e.g., *bil-* 'know how to do, be able', 7) pretence predicates, e.g., *hayal et-* 'imagine' (Csató 2010: 67).

In Johanson's (2013) view, participial complementation markers in Turkic (e.g., Turkish *-DIK*) have a distinctive semantic value: the participles do not refer directly to events as such, but explicitly refer to some potential knowledge about events. The verbal nouns (e.g., Turkish *-mA*), by contrast, are neutral with respect to this semantic value and just refer "to the action, leaving the further interpretation open" (Johanson 2013: 82). Thus, the participial markers convey "the content of cognition", which in turn is "mentally processed". In this connection, Johanson speaks of "knowledge representation", arguing that the participial markers express the notion of "knowledge of a possible fact" and pointing out that this concept is grammaticalized in most Turkic languages (Johanson 2013). Turkish and Noghay, as we have seen, provide morphemes allowing this knowledge-related differentiation. Johanson further raises the theoretical question of whether it is the matrix predicate that selects the subjunctor or vice versa. In his

analysis, the bound morphemes in Turkic languages, which are semantically specific, select the matrix predicates compatible with the notion they express. By comparison, languages like English or German use undifferentiated complementizers whose interpretation mainly depends on the semantics of the matrix predicates. Thus, the situation in Turkic differs from that in languages such as English, German or Russian. For this reason, predicates that usually combine with a verbal noun can also be used with a participle if the semantic notion of 'knowledge of a possible fact' is intended (Johanson 2013).

When a matrix predicate can appear with both complementation markers, the complement clause often implies different readings. The *-DIK* clause, which expresses the notion of 'knowledge of a possible fact', indicates the reading 'that something happens or has happened', whereas the *-mA* clause denotes 'that something will, might happen' (Johanson 2013: 83); consider the examples in (12) and (13) above. In other cases, matrix verbs can get different readings with *-DIK* and *-mA* clauses. In example (14a) the verb *söyle-* expresses 'practical manipulation', whereas in (14b), combined with *-DIK*, it is an utterance predicate meaning 'to say'.

Turkish

(14) a. *Ses* *on-a* [*ağz-ın-ı* *aç-ıp,* *avaz-ı*
voice DEI-DAT mouth-POSS.3-ACC open-CVB voice-POSS.3
çık-tığ-ı *kadar* *bağır-ma-sın-ı*] *söyle-di.*
come_out-DIK-POSS.3 as_much_as shout-MA-POSS.3-ACC say-PST.3
'The voice told him to open (said that he should open) his mouth and (to) shout at the top of his voice.'
(Şafak 1999: 19)

b. *Ses* *on-a* [*ağz-ın-ı* *aç-ıp,* *avaz-ı*
voice DEI-DAT mouth-POSS.3-ACC open-CVB voice-POSS.3
çık-tığ-ı *kadar* *bağır-dığ-ın-ı*] *söyle-di.*
come_out-DIK-POSS.3 as_much_as shout-DIK-POSS.3-ACC say-PST.3
'The voice said to him that he opened/ was opening/ had opened his mouth and shouted/ was shouting/ had shouted at the top of his voice.'

In Noghay, combinational patterns of *-(U)w* and *-GAn* with semantic types of matrix predicates seem to conform to those in Turkish.[14] Thus, complements based on *-GAn* are typically compatible with utterance predicates such as *ayt-* 'say, tell' *sor-* 'ask', *xabarlas-* 'speak'; predicates of knowledge, acquisition and loss of knowledge, immediate perception such as *bil-* 'know', *aŋla-* 'understand', *kör-* 'see', *esit-* 'hear', *tïŋla-* 'listen', *körset-* 'show', *sez-* 'perceive', *esle-* 'notice,

14 The results concerning the combinability of *-GAn* and *-(U)w* with matrix predicates were obtained through a manual scanning of Ajbazova (1996) (see primary sources).

pay attention'; 'predications of mental perception' such as *ökin-* 'regret', *šüšli bol-* 'feel guilty', *šükir et-* 'thank', *säspekle-* 'be confused' etc. Example (15) illustrates a *-GAn*-based complement clause headed by the knowledge predicate *bil-* 'know':

Noghay

(15) *[...]* *[ne* *zat* *üšin* *yeti* *klass-tï* *ḳutïl-yan-da* *maktaw*
what thing for seven class-ACC complete-GAN-LOC distinguished
gramota *em* *sawya-ya* *fotoapparat* *al-yan-ï-n]* *da*
certificate and present-DAT camera receive-GAN-POSS.3-ACC PTC
ärüw *bil-etayan* *edi.*
well know-HAB PST.COP
'She knew very well [...] why she received a certificate of distinction and a camera as a present when she completed the seventh grade.'
(Sikaliev 1968: 8)

Clauses based on *-(U)w* especially occur with volitional predicates such as *süy-* 'want, like, love', *tile-* 'wish', *sayïn-* 'wish', *talapla-* 'request'; phasal predicates such as *ḳoy-* 'cease'; objective modal predicates such as *izle-* 'figure out, estimate'; and the predicate of necessity *tiyisli bol-* 'need'. In example (16), the complement clause *yïlawïn* 'his crying' is headed by the phasal predicate *ḳoy-* 'cease'.

Noghay

(16) *"Öz-iŋ-e* *ḳaldïr-a* *ḳoy* *sol* *aḳ-tï!"* *de-di*
self-POSS.2SG-DAT keep-CVB POVB that salary-ACC say-PST.3
[yïla-w-ï-n] *ḳoy-ïp* *Asantay.*
cry-UW-POSS.3-ACC cease-CVB Asantay.
'"Keep this salary for yourself", said Asantay, ceasing to cry.'
(Ajbazova 1996: 49)

The matrix verbs *bildir-* 'let know', *aŋla-* 'understand', *mut-* 'forget' and *körin-* 'be seen, appear', *yara-* 'suit', *ḳorḳ-* 'be afraid' and *mäne ver-* 'esteem' occur with both markers, allowing the different readings mentioned above. Our general impression from the Noghay data is that some predicates, such as the commentative *süyün-* 'be happy', and the volitional *süy-* 'want, like', can take a complement clause based on the participial marker *-GAn*, even in discourse contexts where, in Turkish, the verbal nominal *-mA* would be more appropriate (17) (Karakoç 2007a [2002]). See also example (5) given above. We cannot elaborate on this observation in this paper as it needs further systematic investigation.

Noghay

(17) *Tete-m* *kör-me-se* *de* *keleyek* *yašaw-um-da,* *o*
grandma-POSS.1SG see-NEG-COND PTC future life-POSS.1SG-LOC she
men-im *üylön-gen-im-i* *köp* *aylak* *süy-etayan* *edi.*
DEI.1SG-GEN marry-GAN-POSS.1SG-ACC very much want-HAB PST.COP
Amma *Alla-dıŋ* *iş-i* *tete-m* *kör-al-ma-dï*
but God-GEN work-POSS.3 grandma-POSS.1SG see-MOD-NEG-PST.3
men *üylön-gen-im-i.*
DEI.1SG marry-GAN-POSS.1SG-ACC
'My grandma, even if she did not see it happen, she wanted me very much to marry in my future life. But, as things happened, she could not see me marry.'
(Karakoç 2005: 227)

Turkish *-DIK* and Noghay *-GAn* do not mark aspectual or temporal notions. Their temporal interpretation depends on the actional content of their verbs, the use of time adverbials, the aspectual/temporal information given in their higher clauses and the discourse constellation: compare (18) and (19), in which *-DIK* is suffixed to the same predicate, *çalış-* 'work, study'. But due to the semantics of the respective matrix predicates, *bil-* 'know' (which can be interpreted flexibly with respect to any placement in time) and *hatırla-* 'remember' (which by way of its lexical semantics typically evokes a past situation), as well as due to the respective discourse constellations (containing other, finite and therefore aspectually specific, verb forms), *çalıştıK-* can be understood as referring to an action that is either ongoing, as in (18), or has taken place in the past, as in (19). In example (20), by contrast, it is the lexical semantics of a finaltransformative embedded verb, *öl-* 'die', that suggests the completedness of the action (Johanson 1971, 2000 for the terminology).

Turkish

(18) *E:vet,* *peki* [*anne-n-in* *ve* *baba-n-ın* *nerede*
yes okay mother-POSS.2SG-GEN and father-POSS.2SG-GEN where
çalış-tığ-ın-ı] *bil-iyor* *mu-sun?* *((1s))* *Nerde* *çalış-ıyor-lar?*
work-DIK-POSS.3-ACC know-PRS Q-2SG where work-PRS-3PL
'Yeah, well, do you know where your Mum and Dad work? ((1s)) Where do they work?'
(Rehbein et al. 2009: EFE07tk_Eni_b_0691_3_SKO)

Turkish

(19) *Ders-ler-im* *de* *pek* *fena* *değil-miş.* *Orta-ymış* *yani.*
subject-PL-POSS.1SG also very bad not-EVID.COP average-EVID.COP that_is
On-dan, *çok* *ders_çalış-ır-dı-m* *amma.* [*Çok* *ders_çalış-tığ-ım-ı*]
DEI-ABL a_lot study-AOR-PST.COP-1SG but a_lot study-DIK-POSS.1SG-ACC
hatırl-ıyor-um.
remember-PRS-1SG
'And my grades weren't so bad at all. They were average, I mean. That's the reason, I did really study hard. I remember that I studied very hard.'
(Rehbein et al. 2009: EFE02tk_Sec_m_0415_1_SKO)

Turkish

(20) *Kaç* *kere* *söyli-ycek-sin* *Samet-cim* [*sakal-lı* *dede-m-in*
how_many time say-YACAK-2SG Samet-DIM beard-DER grandpa-POSS.1SG-GEN
öl-düğ-ün-ü].
die-DIK-POSS.3-ACC
'How many times are you going to say, my dear little Samet, that my bearded grandpa died.'
(Rehbein et al. 2009: EFE08tk_Hai_m_0646_4_SKO)

Ways of expressing aspectual ideas in complement clauses will be discussed in Section 6.4. below.

6.2 Turkish *-(y)Iş*

The verbal noun *-(y)Iş* in Turkish, apart from being a subordinator in complement clauses, is a relatively productive derivational suffix.[15] The Noghay cognate *-(I)s* can only be used derivationally and is not employed in any subordinating function.[16] The status of Turkish *-(y)Iş* as a syntactic complementation marker is, however, far from uncontroversial and is contested for instance by van Schaaik (1999), who analyses it as a derivational nominalizer on the grounds of its reduced argument structure, particularly with respect to direct objects. Although we also found that the inner argument structure of *-(y)Iş*-predicates can be somewhat reduced, we do also find occurrences with direct objects, consider e.g., (22) below.

15 Cf. Banguoğlu ([1990 [1959]), who discusses *-(y)Iş* in his chapters on derivation and on inflection.

16 Erdal (1998: 65) analyses the development of *-Xş*, a productive derivational suffix in Old Turkic, into a complementation marker in southern Turkic languages (such as Turkish, Uzbek and modern Uyghur) as a result of contact with Persian.

In analysing the functions of Turkish *-(y)Iş*, Erdal (1998) takes up the concepts of factivity versus non-factivity. According to him, *-(y)Iş*-complements refer to actual observable facts and to the manner of performing an action (cf. Banguoğlu 1990 [1959]: 422), often serving as topics in their sentences, while *-mA* forms do not. In his view, *-(y)Iş* serves to form "factive imperfective event nominals"; imperfectivity as a key notion correlates with present and also makes possible a reference to manner. In comparison with other morphemes, *-(y)Iş* carries more of a factive implication than for instance *-mA*, sharing the value factivity with *-DIK*. Erdal presents a minimal pair involving *-(y)Iş* and *-DIK*, repeated here as (21a) and (21b), both of which he analyses as factive. But in (21a), the factive event is presented as an ongoing process, whereas in (21b), "the mere fact of coming up the stairs is registered" (Erdal 1998: 60):

Turkish

(21) a. [*Aysel-in* *merdiven-ler-den* *çık-ış-ın-ı*] *gör-dü-m*.[17]
Aysel-GEN stair-PL-ABL come_up-YIS-POSS.3-ACC see-PST-1SG
'I saw Aysel come/ coming up the stairs.'

b. [*Aysel-in* *merdiven-ler-den* *çık-tığ-ın-ı*] *gör-dü-m*.
Aysel-GEN stair-PL-ABL come_up-DIK-POSS.3-ACC see-PST-1SG
'I saw that Aysel came up the stairs.'

Contrary to Erdal's (1998) view, Göksel & Kerslake (2005: 427–428) argue that *-(y)Iş* "cannot be used in sentences that bear upon the factual status of a proposition, nor in those that express a volitional stance (desire, command, request, permission, etc.) towards a situation. It is therefore excluded from most of the contexts in which *-DIK/-(y)AcAK* or *-mA* are used." They also remark: "While the use of *-DIK/-(y)AcAK* expresses the perception of a situation merely as a *fact*, the use of *-mA* or *-(y)Iş* expresses rather the perception of an inherent quality of the event or state".

Returning to Erdal's (1998) examples (21a) and (21b), both *-(y)Iş* and *-DIK* are headed by the 'perception' predicate *gör-* 'see', and both clauses may in principle indicate an 'ongoing' action that takes place simultaneously with the event in the matrix clause. We will argue instead that *-DIK* in (21b) refers to the knowledge of an event that has taken place or is taking place (in the sense of Johanson 2013), whereas *-(y)Iş* in (21a) does not signal this concept and seems to directly refer to the inherent process of an action. In our view *-(y)Iş* does not express factivity or imperfectivity. Rather, a direct reference to the process of an action without reference to the knowledge about it 1) makes the interpretation of an ongoing event more possible for *-(y)Iş* clauses (in the sense of Erdal 1998), 2) often indicates

17 The segmentation, annotation and translation of these cited examples are ours.

countable single instances of events (Göksel & Kerslake 2005: 428; Coşkun 2012: 36–38; van Schaaik 1999), and 3) is often in line with the notion of 'manner'. To sum up, the notions offered in the literature – actual observable facts, ongoing imperfective processes of single events, and manner – seem to be mainly related to the idea that *-(y)Iş* implies a direct reference to the inner process of an action. This is also in accordance with the observation that the actions in question are often presented in complement clauses headed by verbs of perception.[18] Examples (22) and (23), in which several *-(y)Iş*-constructions are employed as subjects of their respective matrix predicates (*anlatıyordu* 'it told', an utterance predicate, and *mühimdi* 'it was important', a predicate expressing a propositional attitude), illustrate how *-(y)Iş* precategorizes the action/ event in terms of manner:

Turkish

(22) [*Bu* *iş-i* *evvelce* *yap-tığ-ı*] *belli:* [*olta-yı*
this work-ACC before make-DIK-POSS.3 obvious fishing rod-ACC
salla-yış-ı], [*tart-ış-ı*]... *bir* *balık* *tut-tu:* [*Balığ-ı*
swing-YIS-POSS.3 pull-YIS-POSS.3 a fish catch-PST.3 fish-ACC
çek-iş-i], [*hiç* *heyecanlan-ma-yış-ı*] *bana* [*bu*
pull-YIS-POSS.3 never get_excited-NEG-YIS-POSS.3 DEI.1SG-DAT this
iş-e *yabancı* *ol-ma-dığ-ın-ı*] *anlat-ıyor-du.*
work-DAT stranger be-NEG-DIK-POSS.3-ACC tell-PRS-PST.COP
'It was obvious that he had done this before. The way he swung and pulled the fishing rod... He caught a fish: The way he pulled up the fish and did not get excited told me that he was no stranger to this task.'
(Abasıyanık 2010: 172)

Turkish

(23) *Bir* *kelime* *ile* *ben-im* *için* *bu* *mesele-ler-in* *kendi-leri* *kadar*
one word with DEI.1SG-GEN for this issue-PL-GEN self-POSS.3PL as
[*onlar-ın* *bana* *gel-iş-leri*], [*ruh_haller-im-i*
they-GEN DEI.1SG-DAT come-YIS-POSS.3PL states_of_mind-POSS.1SG-ACC
benimse-yen *iç-im-de-ki* *yürü-yüş-leri*] *de* *mühim-di.*
accept-PTCP inside-POSS.1SG-LOC-ADJ walk-YIS-POSS.3PL PTC important-PST.COP
'In a word, to me, as important as these matters were themselves, their way of coming towards me and their way of walking inside me, which captured my states of mind, were equally important.'
(Tanpınar 1979: 8)

18 Compare this analysis with that of Boye (2010), who suggests distinguishing between 'knowledge acquired' and 'object of perception' within Cognitive Grammar in terms of a contrast between referring (i.e. referent-stipulating) and nonreferring (i.e. nonreferent-stipulating) status respectively.

Since *-(y)Iş* expresses a concept not signalled by *-DIK* or *-mA*, it occurs in opposition to both, respectively. Compare the minimal pairs in (24a–c) and (25a–c). In (24) *-DIK* and *-(y)Iş*, combined with the knowledge predicate *bil-* 'know', are in semantic opposition, whereas *-mA* does not participate in this opposition (24c). The knowledge predicate *bil-* 'know' does not accord well with *-mA*, which refers to the action as a whole. It is however compatible with *-(y)Iş*. The same can also be observed for perception predicates such as *gör-* 'see'. Note that the clause given in example (21) would also be unacceptable with *-mA*, i.e. **Aysel'in merdivenlerden çıkmasını gördüm*. In (25) the verbal nominal markers *-(y)Iş* and *-mA*, combined with the commentative verb *memnun ol-* 'become glad', are in a semantic contrast, whereas *-DIK* is not possible in the same setting (25c). In these examples, *-mA* directly refers to the action as a whole (25a) whereas in the case of *-(y)Iş* it is the inner process of the action which is referred to (25b). Both in (24b) and (25b), the *-(y)Iş* complements allow for an interpretation of manner. Compare also the minimal pairs containing the commentative predicate *gıcık kap-* 'get irritated' in (26a–b).

Turkish

(24) a. [*Kız-ım-ın* *yemek yap-tığ-ın-ı*] *bil-iyor-um.*
daughter-POSS.1SG-GEN cook-DIK-POSS.3-ACC know-PRS-1SG
'I know that my daughter has cooked (or cooks or is cooking).'
b. [*Kız-ım-ın* *yemek yap-ış-ın-ı*] *bil-iyor-um.*
daughter-POSS.1SG-GEN cook-YIS-POSS.3-ACC know-PRS-1SG
'I know how my daughter cooks.'
c. *[*Kızımın yemek yapmasını*] *biliyorum.*

Turkish

(25) a. [*Kız-ım-ın* *kitap* *oku-ma-sın-dan*] *memnun* *ol-du-m.*
daughter-POSS.1SG-GEN book read-MA-POSS.3-ABL glad become-PST-1SG
'I became happy that my daughter has read (or reads or is reading) books.'
b. [*Kız-ım-ın* *kitap* *oku-yuş-un-dan*] *memnun* *ol-du-m.*
daughter-POSS.1SG-GEN book read-YIS-POSS.3-ABL glad become-PST-1SG
'I became happy how/that my daughter reads books.'
c. *[*Kızımın kitap okuduğundan*] *memnun oldum.*

Turkish

(26) a. *Ayak-lar-ım-ı* *çamur-a* *bas-ıyo-du-m* *böyle.* *Ben*
foot-PL-POSS.1SG-ACC mud-DAT tread_on-PRS-PST.COP-1SG like_this I
gıcık *ol-uyo-du-m.* [*Ayak-lar-ım-ın* *çamur-a*
tickled be-PRS-PST.COP-1SG foot-PL-POSS.1SG-GEN mud-DAT
bas-ma-sın-dan] *gıcık* *k/* *kap-ıyo-du-m* *böyle.*
step-MA-POSS.3-ABL tickled get-PRS-PST.COP-1SG like_this

'I stepped with my feet in the mud, like this. I got tickled. I got t/ tickled from my feet stepping in the mud, like this.'
(Rehbein et al. 2009: EFE08tk_Sed_m_0602_4_SKO)

b. *Ayak-lar-ım-ın* *çamur-a* *bas-ış-ın-dan*
foot-PL-POSS.1SG-GEN mud-DAT tread_on-YIS-POSS.3-ABL
gıcık k/ *kap-ıyo-du-m* *böyle.*
tickled get-PRS-PST.COP-1SG like_this
'I got t/ tickled from the way my feet stepped in the mud, like this.'

In accordance with the semantic value 'direct reference to inner process of an action' *-(y)Iş* prefers to combine with specific kinds of matrix predicates (listed by Noonan 2007 [1985]). In our data, next to predicates of knowledge and immediate perception such as *gör-* 'see', *duy-* 'hear', *fark et-* 'notice, realize, perceive', *gözle-* 'observe', *seyret-* 'watch', *göster-* 'show', see (27), *-(y)Iş* is especially compatible with pretence predicates such as *hayal et-* 'imagine' and commentative predicates such as *gıcık kap-* 'become iritated', *memnun ol-* 'be/become pleased', *üzül-* 'be sorry or worried', *kız-* 'be angry', *sevin-* 'be glad': see (28), as well as (25b) and (26b) above. *-(y)Iş* does not, however, seem to be compatible with volitional, achievement, phasal and objective modal predicates.

Turkish

(27) *Ertesi* *gün,* *köleliğ-im-in* *en* *keyifli* *gün-ü* *ol-du.*
next day slavery-POSS.1SG-GEN SPL delightful day-POSS.3 become-PST.3
Artık *ben-i* *sandalye-ye* *de* *bağla-m-ıyor-du,* *ama* *bütün* *gün-ü*
anymore I-ACC chair-DAT PTC tie-NEG-PRS-PST.COP but whole day-ACC
[*yavaş yavaş* *başka* *bir* *insan* *ol-uş-un-u*] *keyifle,*
slowly slowly another a person become-YIS-POSS.3-ACC enjoying
gözle-mek *için* *karşı-sın-da* *otur-arak* *geçir-di-m.*
observe-MAK for in_front-POSS.3-LOC sit-CVB spend-PST-1SG
'The next day became the most delightful day of my slavery. He did not tie me to the chair anymore, but I spent the whole day sitting in front of him in order to watch with delight how he was slowly becoming a new person.'
(Pamuk 1994: 31)

Turkish

(28) [*Edebiyat-a* *bun-ca* *yakın* *ol-uş-un-a*] *sevin-miş-ti-m.*
literature-DAT this-EQU close be-YIS-POSS.3-DAT be_happy-POST-PST.COP-1SG
'I had felt happy to see how he had such a close relationship with literature (lit.: was so close to literature).'
(Öz 1986: 15)

-(y)Iş can often be found in contexts with iterative time adverbs such as *zaman zaman* 'from time to time, at times', *hergün* 'every day', *her zaman* 'always',

sabahları 'mornings'. *-(y)Iş* also combines with adverbs of time expressing specific moments, such as *o an* 'at just that time', *tam o sırada* 'at just that moment', *birinci kez* 'the first time' etc. In (29) we can observe the use of *beşinci kere* 'the fifth time' as an adverb modifying the complement predication in the construction *Lütfen,* • [*size* • *dört, beşinci kere atışım bu faksı*] 'Please, • it's the • fourth, fifth time I am sending you this fax'; the part in square brackets functions as the nominal predicate of a copular clause; cf. Erdal (1998); Göksel & Kerslake (2005: 428).

Turkish

(29) *Lütfen,* • [*siz-e* • *dört, beşinci kere at-ış-ım buu/* •• *bu faks-ı*].
please DEI.2PL-DAT four fifth time throw-YIS-POSS.1SG DEI DEI fax-ACC
'Please, this is the • fourth, fifth time I'm sending you this • • this fax.'
(Rehbein et al. 2009: EFE08tk_Mek_m_1085_SKO)

Our impression from the data is *-(y)Iş*-clauses are rarely negated. This might be due to the fact that negation is not well compatible with manner descriptions. In other words, one usually does not describe how an action does not take place. In example (30) the negated *-(y)Iş* in a commentative construction does not necessarily imply manner:

Turkish

(30) *Emlakçı-nın görüntü-sü nasıl-dır şimdi? Bodrum-un*
real_estate_agent-GEN appearance-POSS.3 how-COP now basement-GEN
karanlığ-ın-a doğru yol_al-ır-ken, [tombul parmak-lar-ı-yla
darkness-POSS.3-DAT towards proceed-AOR-CVB.COB plump finger-PL-POSS.3-INS
tutun-acak bir yer ol-ma-yış-ın-ın] kızgınlığ-ın-ı da
hold-PTCP one place be-NEG-YIS-POSS.3-GEN anger-POSS.3-ACC also
yüz-ün-de götür-üyor-dur.
face-POSS.3-LOC carry_away-PRS-COP
'What does the real estate agent look like now? In all likelihood, while proceeding towards the darkness of the basement, he carries away on his face the anger at there being no place to grab hold of with his plump fingers (the anger at how there is no place to grab hold of with his plump fingers).'
(Çelik 2011: 108)

6.3 Turkish *-(y)AcAK* and Noghay *-(A)yAK* and *-(A)r eken*

The participial morphemes *-(y)AcAK* in Turkish and its cognate *-(A)yAK* in Noghay express modal prospectivity both in finite clauses and in non-finite clauses, yet with different sets of personal suffixes. The modal prospective markers *-(y)AcAK*

and *-(A)yAK* often denote the planned character of intentions, suggestions etc., related to the future. That a plan held by an animate subject referent is present at the temporal orientation point (reference time) seems to be a crucial motivation for the use of these forms in many cases. In cases where the subject of the sentence is inanimate, these forms convey that there are indications that the prerequisites for the event have been achieved. This allows the speaker to expect/guess that the event is going to happen (Karakoç, forthcoming). Consider the following examples (31–33).

Turkish

(31) *Sor-ar* *sor-maz* [*pişman_ol-acağ-ın-ı*] *san-dı-m,* *ama* *hâlâ*
ask-AOR ask-NEG.AOR regret-YACAK-POSS.3-ACC think-PST-1SG but still
çocuksu *bir* *merak-la* *bana* *bak-ıyor-du.*
childlike a curiosity-INS DEI.1SG-DAT look-PRS-PST.COP
'I thought that as soon as he asked he would regret it, yet he continued to look at me with a childlike curiosity.'
(Pamuk 1994: 37)

Noghay

(32) *Men* *o-ya* *Rauf* *de-gen* *at* *ber-eyeg-im-di* *de,*
I DEI-DAT Rauf say-PTCP name give-AYAK-POSS.1SG-ACC PTC
tuḳïm-ï *Ḳaramov* *bol-ayay-ï-n* *entte* *bil-gen*
family_name-POSS.3 Karamov be-AYAK-POSS.3-ACC in_no_way know-POST
yoḳ *edi-m.*
nonexistent PST.COP-1SG
'I did not know in any way either that I would give him the name Rauf and that his surname would be Karamov.'
(Kapaev 1989: 9)

Noghay

(33) *Dželaldin /.../* [*tez-tez* *üyken* *bol-ïp* *ös-eyeg-in-e*], *em*
Dželaldin fast big become-CVB grow_up-AYAK-POSS.3SG-DAT and
[*sol* *zaman-da* *ata-sï-nïŋ* *öš-i-n*
that time-LOC father-POSS.3-GEN revenge-POSS.3-ACC
al-ayay-ïn-a] *mïradlan-dï.*
get-AYAK-POSS.3-DAT wish-PST.3
'Dželaldin wished to grow up as fast as possible and then to get revenge for his father.'
(Ajbazova 1996: 85)

The modal prospective markers *-(y)AcAK* and *-(A)yAK* are semantically compatible with the same matrix predicates that combine with *-DIK* and *-GAn*, i.e. utterance predicates, predicates expressing propositional attitudes, knowledge

predicates and predications of mental perception (examples 31–32). They are semantically not compatible with predicates of practical manipulation, such as 'to order, command', 'to oblige', 'to prevent', 'to allow', 'to advise', nor with volitional predicates such as 'to want', 'to request', 'to desire'. However, the volitional predicates *um-* 'to hope', *ümit et-* 'to hope' in Turkish and *mïradlan-* 'to wish' in Noghay do combine with these markers (example 33).

In Turkish, commentative predicates, such as *memnun ol-* 'be pleased', *sevin-* 'be glad', *üzül-* 'be sorry', *şaşır-* 'be surprised', *kork-* 'be afraid, fear' can head clauses based on *-(y)AcAK*, see (34). We observe however that these commentative predicates, which usually are compatible with the verbal nominal marker *-mA*, often prefer combinations with the complex complementation marker *-(y)AcAk ol-ma-* instead of the simple marker *-(y)AcAK*.

Turkish

(34) *On_beş* *gün* *boyunca* *okul-a* *hiç* *gi/git-me-di-n.* *Şimdi*
fifteen day throughout school-DAT at_all go-NEG-PST-2SG now
[*okul-un* *aç-ıl-cağ-ın-a*] *sevin-m-iyo* *mu-sun?*
school-GEN open-PASS-YACAK-POSS.3-DAT be_happy-NEG-PRS Q-2SG
'You haven't been to school for a fortnight. Are you happy now that your school will open again?'
(Rehbein et al. 2009: EFE08tk_Sed_m_0602_3_SKO)

The modal prospective in *-(A)r* in Noghay typically indicates intentions, promises, expectations, guesses, etc. that are not based on plans and just expressed spontaneously in the given speech situation. This modal prospective form in combination with the copular marker *eken* can rarely be found in complement clauses (Karakoç, forthcoming): see (35). The Turkish cognate *-(V)r* does not occur in complement clauses.

Noghay

(35) *Sonï* *man,* *gražđan-lar,* *siz-ge* *sol* *ķader* *yaxšïlïk* *et-etayan*
here with citizen-PL DEI.2PL-DAT that much goodness do-PTCP
velikiy *Germaniya* *ald-ïn-da* [[*siz* *öz* *boriš-ïŋïz-dï*
great Germany front-POSS.3-LOC DEI.2PL self debt-POSS.2PL-ACC
aķ *yüreg-iŋiz* *ben* *toltïr-ar* *eken-iŋiz-ge*] *bek*
white heart-POSS.2PL with pay-AR EKEN-POSS.2PL-DAT very
sen-etayan-ïm-dï] *ayt-ïp,* *men* *söz-im-di* *ķutïl-a-man.*
hope-HAB-POSS.1SG-ACC say-CVB DEI1SG word-POSS.1SG-ACC finish-PRS-1SG
'Thus, my citizens, I finish my words by saying that I hope that you will pay your debts wholeheartedly to the great Germany, which has been so kind to you.'
(Karakoç 2007a [2002]: 354)

6.4 Complex morphemes expressing aspectual notions

Complex verbal forms in complement clauses are employed to express aspectual notions, which are otherwise left unexpressed in the non-finite forms. In Turkic languages, aspect is a grammatical category expressed in both finite and non-finite clauses, albeit to different extents. Without any modification of the actional content of a lexeme, an action can be considered from different points of view, relating to its initial stage, its course, and its end stage (Johanson e.g., 1971, 2000). Aspect is a question of different subjective perspectives on one and the same action. Two types of subjective points of view in Turkic are "intraterminality" [+INTRA] and "postterminality" [+POST]. "Intraterminality" is the perspective of a point of view that does not show the borders of an action but rather hides them. It only implies the view of the action after its beginning and before its ending, independently of its actional content (Johanson 2000: 76–102). "Postterminality" views an event at the point where the critical border of the action has already been exceeded (Johanson 2000: 102–135). Turkish and Noghay have sets of aspectual forms that differ between finite and non-finite clauses to a certain extent.

In Turkish and Noghay complement clauses, aspectual ideas can be expressed by means of morphemically complex forms. In Noghay, intraterminality is denoted by means of complex participial forms, as well as by copulative complementation markers which contain the copular morpheme *eken*. The morphemically complex participles in *-(A)yatḳan*, *-(A)yatïryan*, *-Atayan* have developed from postverbial constructions as a result of present renewals (see Johanson 1995; for the structure of these postverbial participles, see Karakoç 2005, 2007a [2002], 2007b). The complex form *-AdI eken* consists of the simple present marker *-AdI* and the static copular marker *eken*, which carries the nominal morphology necessary for complementation. In order to be able to describe differences between these forms, we need to introduce the notion of "focality" as a further dimension. "Focality" is a scalar notion and concerns a more or less focused psychological interest in the situation. Thus, intraterminal and postterminal items may each exhibit higher or lower degrees of focality (Johanson 2000: 38). Accordingly, the intraterminal participles *-(A)yatḳan* and *-(A)yatïryan*, which are the non-finite counterparts of the finite intraterminal form *-(A)yatïr*, express high-focal intraterminality and correspond to progressive forms, see examples (36) and (37):

Noghay

(36) /.../ [*ḳol-ïn-da* *tayay-ï* *bol-yan,* *bir* *kisi-diŋ*
hand-POSS.3-LOC stick-POSS.3 be-PTCP a person-GEN
kel-eyatkan-ï-n] *kör-di-m.*
come-AYATKAN-POSS.3-ACC see-PST-1SG
'I saw that a person holding a stick in his hand was just coming.'
(Džanbidaeva & Ogurlieva 1995: 240)

Noghay

(37) *Asantay-dïŋ* *yüreg-in-e* *buz* *ḳuy-ïl-dï* *em* *bu,* [*bir* *bäle*
Asantay-GEN heart-POSS.3-DAT ice put-PASS-PST.3 and he a threat
yuwïḳla-p *kel-eyatïryan-ï-n*] *sez-ip,*
approach-CVB come-AYATIRGAN-POSS.3-ACC sense-CVB
kirpik-ler-i-n *ḳay-ïp* *basla-dï.*
eye_lid-PL-POSS.3-ACC blink-CVB begin-PST.3
'Asantay suddenly felt ice-cold in his chest, and he sensed that a threat was approaching and began to blink his eye lids.'
(Ajbazova 1996: 34)

The participial form *-Atayan*, occurring in both finite and different non-finite syntactic constructions (yet with different sets of subject markers), expresses habituality (38):

Noghay

(38) *Yemis* *bil-me-y* *edi* [*Medey-diŋ* *yatlaw-lar* *yaz-atayan-ï-n*],
Yemis know-NEG-CVB PST.COP Medey-GEN poem-PL write-HAB-POSS.3-ACC
bil-me-y *edi* [*bala-sï* *üy-de* *oḳ-uw-ï-n*
know-NEG-PRS PST.COP child-POSS.3 home-LOC read-UW-POSS.3-ACC
et-pe-y *škola-ya* *bar-atayan-ï-n*].
do-NEG-CVB school-DAT go-HAB-POSS.3-ACC
'Yemis did not know that Medey used to write poetry and he did not know that his child used to go to school without doing his homework.'
(Kazakov 1983: 31)

The non-focal form *-AdI eken*, which consists of *-AdI* and the copular marker *eken*, denotes a more general intraterminal view without a sharp focus on the ongoing situation (39). The first part (*-AdI*) signals non-focal and general present/future in finite clauses. The second part *eken* plays an essential role in carrying the morphology (possessive + case) necessary in the complementation process. In this context, it is important to mention that the copular form *eken* does not imply

any evidential meaning in embedded clauses, although it occurs as an evidential marker in finite clauses (Karakoç 2005, 2007a [2002]).[19]

Noghay

(39) *Ramazan tïs-ḳa šïḳ-ḳanlayï oḳ [at-tïŋ art-ïn-da-yï*
Ramazan outside-DAT go_out-CVB just horse-GEN back-POSS.3-LOC-ADJ
oŋ ayay-ï aḳsa-y-dï eken-i-n] közle-p al-dï.
right foot-POSS.3 limp-ADI EKEN-POSS.3-ACC notice-CVB POVB-PST.3
'Ramazan, as soon as he went outside, immediately noticed that the horse was limping on its right hind foot.'
(Kurmangulova et al. 1991: 5)

In Turkish, there are no complex (postverbial) participials at all and the occurrence of copulative complementation markers expressing intraterminality is much rarer than in Noghay. In order to present an event as ongoing (high-focal) the following copulative complementation markers can be used in Turkish: *-(Ø)Iyor ol-duK-*, *-(Ø)Iyor ol-ma-*, *-(Ø)Iyor ol-uş-* or *-mAktA ol-duK-*.[20] Like copulative complementation markers in Noghay, these Turkish forms consist of two parts. The first part, *-(Ø)Iyor/ -mAktA*, expresses the 'present' in finite clauses. The second part, *ol-duK-*, *ol-ma-* or *ol-uş-*, consists of the copular verb *ol-* 'become/ be' + *-DIK*-participle, the *-mA*-verbal noun or the *-(y)Iş*-verbal noun, respectively. In these complex structures, neither the Noghay copular form *eken* nor the Turkish copular verb *ol-* implies any content-related information, such as evidentiality, presumption etc. The only function of these copular markers is to perform the syntactic embedding. See examples (40) to (42). The complex forms used in examples (40–41) are *-(Ø)Iyor ol-duk-* and *-(Ø)Iyor ol-uş-* respectively, while example (42) illustrates the use of *-mAktA ol-duk-*. See also example (8) above, in which the complex intraterminal form *-(Ø)Iyor ol-ma-* is used.

19 It seems safe to say that evidentiality is one of the semantic distinctions that gets lost in deranked constructions such as complement clauses.

20 The use of *ol-ma-* and *ol-uş-* in combination with the finite form *-mAktA* is rather unusual.

Turkish

(40) *Ben-i* *sev-en-ler-in* *sık_sık* *ben-i* *düşün-üp,* [[*İstanbul-un* *bir*
I-ACC love-PTCP-PL-GEN often DEI.1SG-ACC think-CVB Istanbul-GEN a
köşe-sin-de *aptalca* *bir* *meşgale-yle* *hâlâ* *oyalan-ıyor*
corner-POSS.3-LOC footling a dealing-with still hang_out-IYOR
ol-duğ-um-u], *hatta* [*başka* *bir* *kadın-ın* *peş-in-den*
be-DUK-POSS.1SG-ACC even another a woman-GEN behind-POSS.3-ABL
git-tiğ-im-i] *hayal_et-me-leri*] *huzursuz* *ruh-um-a*
go-DIK-POSS.1SG-ACC imagine-MA-POSS.3PL peaceless soul-POSS.1SG-DAT
büsbütün *azap_ver-iyor.*
completely torment-PRS.3
'It completely torments my peaceless soul that those who love me think of me often and imagine me engaged in footling dealings somewhere in Istanbul, or even going after another woman.'
(Pamuk 1998: 14)

Turkish

(41) *Bu* *kilise* *hizmetkârı-nın* *iki_büklüm* *beden-i* *kendin-den* *çok* *on-u*
this church servant-GEN bent body-POSS.3 self-ABL much DEI-ACC
gör-en-ler-e *acı* *ver-iyor-du.* *Öte* *yan-dan* [*iş-in-i*
see-PTCP-PL-DAT pain give-PRS-PST.COP other side-ABL work-POSS.3-ACC
tutku-yla *yap-ıyor* *ol-uş-u*] *insan-lar-da-ki* *ölüm* *beklentisin-i*
passion-INS make-IYOR be-YIS-POSS.3 man-PL-LOC-ADJ death expectation-ACC
ertel-iyor-du.
postpone-PRS-PST.COP
'The bent body of this church servant gave more pain to those seeing him than it did to himself. On the other hand, that he was doing his work with passion postponed the expectation of death in people.'
(Çelik 2011: 9)

Turkish

(42) *Baş-ım-ı* *çevir-ince* [*onlar-ın* *bayır-ın* *başka* *taraf-ın-dan*
head-POSS.1SG-ACC turn-CVB they-GEN slope-GEN other side-POSS.3-ABL
yavaş *yavaş* *in-mekte* *ol-duk-ların-ı*] *gör-me-yeyim* *mi?*
slowly slowly come_down-MAKTA be-DIK-POSS.3PL-ACC see-NEG-OPT.1SG Q
'Turning my head, I realized to my surprise that they were slowly descending the slope from the other side.'
(Güntekin 2010: 35)

Table (4) summarizes Turkish and Noghay intraterminal forms in complement clauses and their counterparts in finite clauses.

Table 4: Intraterminal forms in complement clauses and their counterparts in finite clauses

	Intraterminals in complement clauses	**Intraterminals in finite clauses**
[nog]	*-(A)yatkan / -(A)yatïrγan*	*-(A)yatïr*
	-Atayan	*-Atayan*
	-AdI eken	*-AdI*
[tur]	*-mAktA olduK-*	*-mAktA*
	-(ø)Iyor olduK- / olma- / oluş-	*-(ø)Iyor*

Postterminality in Noghay finite clauses is expressed by *-GAn*. In complement clauses, this aspectual idea is denoted by the complex form *-GAn eken*: see (43). The situation seems to be a bit tricky in Noghay, since the participial morpheme *-GAn* occurs with different semantic contents in finite versus non-finite clauses. In finite clauses, it is a purely postterminal marker. In complement clauses, however, it does not carry any aspectual value of its own. Without the copular marker *eken*, *-GAn* only expresses the concept 'content of knowledge', as discussed in Section 6.1. In complement clauses, *-GAn* needs to be accompanied by the copular marker *eken* in order to express postterminality.

Noghay

(43) *Bala-lar-dïŋ* *ana-sï* *da* *razï_bol-a-dï* *bala-lar-dï* *ber-ip*
child-PL-GEN mother-POSS.3 PTC agree-PRS-3 child-PL-ACC give-CVB
yiber-mege. Ol *bil-e-di* [*öz-ü-nüŋ* *sonγï* *kün-lör-ü*
send-MAGA she know-PRS-3SG self- POSS.3-GEN last day-PL-POSS.3
ķal-yan *eken-i-n*].
remain-GAN EKEN-POSS.3-ACC
'The mother of the children also agrees to send her children away. She knows that these are the last days of her life.'
(Karakoç 2005: 224)

Postterminality in Turkish finite clauses is expressed by the verbal suffix *-mIş*. *-mIş* also functions as an evidential marker in finite clauses. In complement clauses, it does not occur on its own, but needs to be accompanied by the copular forms *ol-duK-* or *ol-ma-*. Thus, postterminality is denoted by the complex morpheme *-mIş ol-duK* or *-mIş ol-ma-*. The *-mIş* part of this structure explicitly expresses anteriority in the meaning of 'having done'. It is important to point out that *-mIş*, expressing evidentiality in finite clauses, does not imply any evidential meanings when occurring in combination with *ol-duK-/ol-ma-* in embedded clauses. Example (44) illustrates the use of *-mIş ol-duK*:

Turkish

(44) *Mimar-ın* *veya* *hayrat* *sahibi-nin* *dik-tiğ-i* *ağac-ın*
architect-GEN or charity owner-GEN plant-DIK-POSS.3 tree-GEN
büyü-düğ-ün-ü *gör-üp* *gör-me-me-si-nin* *ehemmiyet-i*
grow_out-DIK-POSS.3-ACC see-CVB see-NEG-MA-POSS.3-GEN importance-POSS.3
yok-tu. [[*Dikil-miş* *ol-duğ-un-u*] *bil-me-si*]
nonexistent-PST.COP plant-MIS be-DIK-POSS.3-ACC know-MA-POSS.3
yeter-di.
sufficient-PST.COP
'It was not important for the architect or the charity owner to see whether the tree(s) he planted grew up. It was sufficient for him/her to know that it had been planted.'
(Tanpınar 1979: 55)

Table (5) summarizes the Turkish and Noghay postterminal forms in complement clauses and their counterparts in finite clauses.

Table 5: Postterminal forms in complement clauses and their counterparts in finite clauses

	Postterminals in complement clauses	**Postterminals in finite clauses**
[nog]	*-GAn eken*	*-GAn*
[tur]	*-mIş olduK / -mIş olma-*	*-mIş*

As previously described, tense, evidentiality or other modal values such as 'presumption' cannot be expressed in complement clauses by means of the morphologically bound complementation markers. The means to express these semantic notions are mainly restricted to clauses with finite morphology; in non-finite constructions, such ideas can only be expressed by adverbial means. The only semantic idea that is grammatically expressed in nominalized complement clauses is aspect.

7 Bound complementation markers in non-complementizing functions

Complementizing suffixes in Turkish and Noghay are polyfunctional. In principle, they can be used as relativizers, as parts of adverbializers, as finite forms, and as derivational forms. Their range of functions is similar, but not entirely identical in Turkish and Noghay. Table (6) gives an overview of the other possible uses of bound complementation markers.

Table 6: Other occurrences of bound complementation markers

		Complementation markers	in attributive clauses	in adverbial clauses	in finite clauses	as derivational forms
Verbal-nominal morphemes	[tur]	*-mAK*	–	+	–	+
		-mA	–	+	–	+
		-(y)Iş	–	+	–	+
	[nog]	*-(U)w*	–	–	–	+
		-MAGA	–	–	–	–
Participial morphemes	[tur]	*-DIK*	+	+	–	+
		-(y)AcAK	+	+	+	+
	[nog]	*-GAn*	+	+	+	+
		-(A)yAK	+	+	+	+
		-Atayan	+	+	+	–
		-(A)yatḳan	+	+	+[21]	–
		-(A)yatïryan	+	+	+	–

Based on this table, the following remarks can be made: 1) In the verbal-nominal markers, the derivational function is primary: their complementizing function is a later development. 2) In the participial markers, the subordinating function is primary, whereas their derivational function is a later development. The postverbial participles, which historically represent a later development in Noghay, do not appear in derivational functions. 3) The participial markers can also occur in attributive clauses and, apart from Turkish *-DIK*, can appear in finite clauses, yet with different sets of personal suffixes. 4) Both the verbal-nominal and the participial markers are found in the structures of complex subjunctors used in adverbial clauses. In such complex subjunctors they are often combined with possessive suffixes, case markers, and/or postpositions. The forms and functions of these complex subjunctors differ in Turkish and Noghay.

In this section, we show some other syntactic uses of the complementation markers. Example (45) illustrates the use of the Noghay participial form *-GAn* as a relativizer in an attributive clause, while example (46) shows the use of the same

21 The Noghay participles *-Ayatḳan* and *-Ayatïryan* can also occur as finite forms, but only in combinations with copular markers and in special discourse types (Karakoç 2005).

morpheme as a finite morpheme in a finite clause. Note that they have different sets of personal suffixes depending on whether they are employed in finite or non-finite clauses; example (46) does not fully illustrate this phenomenon, being a third-person form.

Noghay

(45) *Üy-de* *oltïr-yan* *ḳart* *tuwïl* *eken,* *a* *mazallï* *yigit* *eken.*
house-LOC sit-GAN old_man not EVID.COP but strong young_man EVID.COP
'The person who was sitting at home was (as it turned out) not an old man, but a strong young man.'
(Kazakov 1983: 10)

Noghay

(46) *Burïn-yï* *noyay-lar-da* *yaŋï* *yïl* *mart* *ayïn-nan* *basla-n-yan.*
early-ADJ Noghay-PL-LOC new year march month-ABL begin-PASS-GAN
Noyay-lar *mart* *ayïn-a* *nawruz* *ayï* *de-p* *ayt-ḳan-lar.*
Noghay-PL march month-DAT Nawruz month-POSS.3 saying say-GAN-3PL
'With the early Noghays, the New Year began in March. The Noghays called the month of March the month of Nawruz.'
(Džanbidaeva & Ogurlieva 1995: 10)

Example (47) illustrates the use of the Turkish complex adverbializing morpheme *-DIKçA*, which consists of *-DIK* and an equative suffix:

Turkish

(47) *Ha,* *ya* *balkon-a,* *ya* *cam-ın* *ön-ün-e* *koy-ar-dı-k.*
IJ or balcony-DAT or pane-GEN front-POSS.3-DAT put-AOR-PST.COP-1PL
[*Rüzgar* *vur-duk-ça*] *böyle* *"Pıuuuır"* *dön-er-di* *bu.*
wind beat-DIK-EQU like_this ONO turn-AOR-PST.COP this
'Yeah, we used to put it either on the balcony, or in front of the windowpane. The more the wind beat [against it], the more it would turn, this one, making a pirrrrr sound.'
(Rehbein 2009: Selbtk_Guz_m_0454ab_ENF)

There has been extensive discussion on the boundaries between derivational and inflectional morphology in the Turkic context (see Section 6.3). Turkish and Noghay complementizing suffixes can also occur as derivational suffixes. The distinction between their derivational and their complementizing use crucially depends on the unfolding of an argument structure and on the use of adverbs versus adjectives. In example (48) below two different uses of the Turkish suffix *-(y)Iş* are illustrated. The suffix in *ordaki bakış yerleri* 'the points of view there' is derivational: *Bakış* forms a compound noun together with *yerleri*, and the modifying expression *ordaki* is an adjectival form derived from a locative expression

(*orda* 'there'). The second occurrence, *o insanın ayakta duruşu* 'how this person is standing on his feet' is clausal. *Duruş* is modified not by an adjective, but by an adverbial expression, which is a locative form: *ayakta* 'on his feet'.

Turkish

(48) *( ) yani önemli ol-an o-nun ayağ-ın-da-ki*
that_is important be-PTCP DEI-GEN foot-POSS.3-LOC-ADJ
değil, önemli ol-an orda-ki bak-ış yer-ler-i. • [O
not important be-PTCP there-ADJ look-YIS place-PL-POSS.3 DEI
insan-ın ayak-ta dur-uş-u], yalnız ayak-lar-ın gör-ün-me-si.
person-GEN foot-LOC stand-YIS-POSS.3 only foot-PL-GEN see-PASS-MA-POSS.3
'( ) I mean, what's important is not what's on his foot, what's important is the points of view there. • How this person is standing on his feet, that only the feet are seen.'
(Rehbein 2009: EFE05tk_Can_m_0395_2_ENF)

8 Summary

The aim of this paper has been to discuss Turkish and Noghay complement clauses and complementation markers from a semantic perspective. Starting from the typological fact that complementation markers in Turkic languages are suffixes forming participles and verbal nouns, we illustrated their morphosyntactic integration into the superordinate structure, showing how complementation takes place at the level of aspect and illocution, with some repercussions on how subordinated propositions are semantically presented. We conducted our analysis with a form-by-form approach, operationalizing semantics in terms of knowledge-related and aspectual-temporal meanings, as well as in terms of interactions with matrix predicates.

We can sum up that the following concepts/notions are expressed grammatically in Turkish and Noghay: the participial forms *-DIK* and *-GAn* express the concept 'knowledge of a possible fact' (in the sense of Johanson 2013). The other participial forms, Turkish *-(y)AcAK* and Noghay *-(A)yAK* and *-(A)r eken*, while expressing the same notion, also convey a notion of modal prospectivity in the sense of intentions, suggestions, proposals etc. As for the verbal nouns, *-mA* and *-(U)w* directly refer to an action as a whole, while the Turkish form *-(y)Iş* adds a further distinction by directly referring to the inherent process of an action. In both languages, but to different extents, there exist morphologically complex forms, containing copular elements or postverbial constructions, which express aspectual distinctions in complement clauses. In this connection, Turkish *-(ø)Iyor olduK-*, *-(ø)Iyor olma-*, *-(ø)Iyor oluş-*, *-mAktA olduK-* and Noghay

-AdI eken, *-Atayan*, *-(A)yatḳan*, *-(A)yatïryan* express intraterminal ideas including, among others, the distinctive notion of focality. The other complex forms, Turkish *-mIş olduK-*, *-mIş olma-* and Noghay *-GAn eken*, express the aspectual notion of postterminality.

Due to syntactic deranking, there are some categorial restrictions that affect the semantics of the constructions in context. For instance, temporal interpretation in complement clauses depends on information given as part of the actional content of the verb, time adverbials, the higher clause or the discourse constellation. Another observation concerns the realization of the semantic category of evidentiality/ presumption: the Noghay copular form *eken*, an evidential marker in finite clauses, does not have any evidential meaning in complement clauses.

Acknowledgment

We would like to thank Kasper Boye and Petar Kehayov, the editors of the present volume, as well as an anonymous reviewer for very thoroughly and critically reading and commenting on a first version. Additional thanks go to Jochen Rehbein for discussing the complementizer status of the Turkish suffixes and for his critical and helpful comments on an early version. Additional thanks go to Everett Thiele for proofreading our English.

HIAT conventions used in the spoken language examples

Hm̌	falling-rising tone	/	repair
	(on aspirated nasal)	(( ))	comment within a tier
Hm̄	level tone	( )	incomprehensible
Hm̂	rising-falling tone	•	pause of short duration
Hḿ	rising tone	• •	pause of medium duration
Hm̀	falling tone	• • •	pause of long duration
.	utterance-final sign		(shorter than a second)
	after an interjection (Hm̀)	((2,5s))	pause longer than a second
...	abortion of an utterance		

Primary sources

Abasıyanık, Sait Faik. 2010. *Son kuşlar* [The last birds]. İstanbul: Yapı Kredi Yayınları.
Ajbazova, E. S. (ed.). 1996. *Nogaj literaturasy. Učebnik-xrestomatija* [Noghay literature: Textbook-Chrestomathy]. Čerkessk.
Çelik, Jaklin. 2011. *Öfkenin şenliği* [The feast of wrath]. 1st edn. İstanbul: İletişim Yayınları.
Džanbidaeva, Dž. K. & E. S. Ogurlieva (eds.). 1995. *Nogay literaturasï. 6. klass üšin* [Noghay literature: For grade 6]. Čerkessk: Karašay-Šerkeš kitap baspasï.
Güntekin, Reşat Nuri. 2010. *Çalıkuşu* [The wren]. İstanbul: İnkılap Kitabevi.
Kapaev, Isa. 1989. *Sülder. Birinši kniga* [Sülder. First Book]. Čerkessk.
Kaygusuz, Sema. 2006. *Yere düşen dualar* [Prayers falling to the ground]. 3rd edn. İstanbul: Doğan Kitapçılık.
Kazakov, Valerij. 1983. *Togïz ḳaptal. Povest'ler men xabarlar* [Nine kaftans: Tales and stories]. Čerkessk.
Kurmangulova, Š. A., N. A. Daulova & M. S. Kireev (eds.). 1991. *Nogaj literaturasy. Xrestomatija* [Noghay literature: Chrestomathy]. Čerkessk.
Öz, Erdal. 1986. *Gülünün solduğu akşam* [The night the rose wilted]. 27th edn. İstanbul: Can Yayınları.
Pamuk, Orhan. 1994. *Beyaz kale* [The white castle]. İstanbul: İletişim Yayınları.
Pamuk, Orhan. 1998. *Benim adım kırmızı* [My name is red]. İstanbul: İletişim Yayınları.
Rehbein, Jochen. 2009. Rehbein-ENDFAS: (Die Entwicklung narrativer Diskursfähigkeiten im Deutschen und Türkischen in Familie und Schule [The development of narrative discourse abilities in German and Turkish in family and school]). Universität Hamburg, EXMARaLDA: SFB 538-Korpora: http://www.exmaralda.org/sfb_e5.html.
Rehbein, Jochen, Annette Herkenrath & Birsel Karakoç. 2009. Rehbein-SKOBI (Sprachliche Konnektivität bei bilingual türkisch-deutsch aufwachsenden Kindern [Linguistic connectivity in bilingual Turkish-German children]). Universität Hamburg, EXMARaLDA: SFB 538-Korpora: http://www.exmaralda.org/sfb_ e5.html.
Sikaliev, Ašim. 1968. *Jyltyn. Rasskazlar em novellalar* [Spark: Tales and short stories]. Stavropol' Knigoizdatel'stvosynyŋ Karašaj-Šerkeš Böligi.
Şafak, Elif. 1999. *Şehrin Aynaları* [The mirrors of the city]. 7th edn. İstanbul: Metis Yayınları.
Tanpınar, Ahmet Hamdi. 1979. *Beş Şehir* [Five cities]. 6th edn. İstanbul: Dergâh Yayınları.
Toptaş, Hasan Ali. 1995. *Gölgesizler* [The shadowless]. 5th edn. İstanbul: Doğan Kitap.

References

Abney, Steven Paul. 1987. *The English noun phrase in its sentential aspects*. Cambridge, MA: MIT unpublished PhD dissertation. http://www.vinartus.com/spa/87a.pdf (accessed 1 July 2013).
Aikhenvald, Alexandra Y. 2006. Complement clause types and complementation strategies in Tariana. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation. A cross-linguistic typology*, 178–203. Oxford: Oxford University Press.
Baker, Mark C. 1986. Syntactic affixation and English gerunds. Mimeo, MIT. In *Proceedings of the West Coast conference on formal linguistics*, Stanford University, Volume IV.

Baker, Mark C. 1988. *Incorporation. A theory of grammatical function changing*. Chicago: The University of Chicago Press.

Baker, Mark C. 2012. Degrees of nominalization: Clause-like constituents in Sakha. In Jaklin Kornfilt & John Whitman (eds.), *Nominalizations in linguistic theory.* [Special issue]. *Lingua* 121(7). 1164–1193.

Banguoğlu, Tahsin. 1990 [1959]. *Türkçenin grameri* [Grammar of Turkish], 3rd edn. Ankara: Türk Dil Kurumu (Türk Dil Kurumu Yayınları [Publications of the Turkish Language Society]: 528).

Bernstein, Judy. 2001. The DP hypothesis: Identifying clausal properties in the nominal domain. In Mark R. Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 536–561. Oxford: Blackwell.

Borsley, Robert D. & Jaklin Kornfilt. 2000. Mixed extended projections. In Robert D. Borsley (ed.), *The nature and function of syntatic categories*. (Syntax and Semantics, Volume 32), 101–131. Bingley, UK: Emerald Group Publishing Limited

Boye, Kasper. 2010. Reference and clausal perception-verb complements. *Linguistics* 48(2). 391–430.

Boye, Kasper & Peter Harder. 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in Language* 31(3). 569–609.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1993 [1981]. *Lectures on government and binding*, 7th edn. Berlin: De Gruyter.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: The MIT Press.

Chomsky, Noam. 2005 [1980]. *Rules and representations*. New York: Columbia Universitiy Press.

Coşkun, Hatice. 2010a. Embedding indirective (evidential) utterances in Turkish. In Gabriele Diewald & Elena Smirnova (eds.), *Linguistic realization of evidentiality in European languages*. Berlin & New York: Walter de Gruyter, 187–221.

Coşkun, Hatice. 2010b. Question elements in Turkish complement clauses. *Turkic Languages* 14. 43–68.

Coşkun, Hatice. 2012. *Komplementsätze im Türkischen* [Complement clauses in Turkish]. Istanbul: Litera yayıncılık.

Cristofaro, Sonia. 2005. *Subordination* (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.

Csató, Éva Á. 1999. Modalität in türkischen Komplementsätzen und ihre Entsprechungen im Deutschen [Modality in Turkish complement clauses and their equivalences in German]. In Lars Johanson & Jochen Rehbein (eds.), *Türkisch und Deutsch im Vergleich* [Turkish and German in comparison], 23–32. Wiesbaden: Harrassowitz.

Csató, Éva Á. 2010. Two types of complement clauses in Turkish. In Hendrik Boeschoten & Julian Rentzsch (ed.), *Turcology in Mainz/ Turkologie in Mainz* (Turcologica 82), 107–122. Wiesbaden: Harrassowitz.

Csató, Éva & Birsel Karakoç. 2006 [1998]. Noghay. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*. 2nd edn., 333–343. London: Routledge.

Deutscher, Guy. 2006. Complement clause types and complementation strategies in Akkadian. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation. A cross-linguistic typology*, 159–177. Oxford: Oxford University Press.

Dixon, R. M. W. 1995. Complement clauses and complementation strategies. In F. R. Palmer (ed.), *Grammar and meaning: essays in honour of Sir John Lyons*, 175–220. Cambridge: Cambridge University Press.

Dixon, R. M. W. 2006. Complement clauses and complementation strategies in typological perspective. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation: A cross-linguistic typology*, 1–48. Oxford: Oxford University Press.

Dixon, R. M. W. 2009. The semantics of clause linking in a typological perspective. In R. M. W. Dixon & Alexandra Aikhenvald (eds.), *The semantics of clause linking: A cross-linguistic typology* (Explorations in Linguistic Typology 5), 1–55. Oxford: Oxford University Press.

Dixon, R. M. W. & Alexandra Y. Aikhenvald (eds.). 2006. *Complementation: A cross-linguistic typology*. Oxford: Oxford University Press.

Ehlich, Konrad. 1994. Funktionale Etymologie [Functional etymology]. In Gisela Brünner & Gabriele Graefen (eds.), *Texte und Diskurse. Methoden und Forschungsergebnisse der Funktionalen Pragmatik* [Texts and discourses: Methods and research results of Functional Pragmatics], 68–82. Opladen: Westdeutscher Verlag.

Ehlich, Konrad & Jochen Rehbein. 1976. Halbinterpretative Arbeitstranskriptionen (HIAT) [Semi-interpretative working transcriptions (HIAT)]. *Linguistische Berichte* [Linguistic Reports] 45. 21–41.

Erdal, Marcel. 1998. On the verbal noun in *-(y)Iş*. In Kâmile İmer & Leyla Subaşı-Uzun (eds.), *Doğan Aksan armağanı* [Festschrift for Doğan Aksan], 53–68. Ankara: Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi.

Erguvanlı Taylan, Eser. 1994. Yantümcelerde tamlayan ekinin işlevi üzerine [On the function of the genitive in subordinate clauses]. *Dilbilim Araştırmaları* [Linguistic Investigations] 1994. 31–40.

Fleck, David W. 2006. Complement clause types and complementation strategies in Matses. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation. A cross-linguistic typology*, 224–244. Oxford: Oxford University Press.

Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A comprehensive grammar* (Routledge Comprehensive Grammars). London: Routledge.

Haspelmath, Martin. 1989. From purposive to infinitive: A universal path of grammaticization. *Folia Linguistica Historica* 10 (1–2). 287–310.

Hellwig, Birgit. 2006. Complement clause types and complementation strategies in Goemai. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation. A cross-linguistic typology*, 204–223. Oxford: Oxford University Press.

Herkenrath, Annette & Birsel Karakoç. 2007. Zur Morphosyntax äußerungsinterner Konnektivität bei mono- und bilingualen türkischen Kindern [On the morphosyntax of utterance-internal connectivity in monolingual and bilingual Turkish children]. In Hendrik Boeschoten & Heidi Stein (eds.), *Einheit und Vielfalt in der türkischen Welt. Materialien der 5. Deutschen Turkologenkonferenz, Universität Mainz, 4. – 7. Oktober 2002* [Unity and diversity in the Turkish world: Materials of the 5th German Conference of Turcologists, University of Mainz, 4–7 October 2002], 131–160. Wiesbaden: Harrassowitz.

Herkenrath, Annette, Birsel Karakoç & Jochen Rehbein. 2003. Interrogative elements as subordinators in Turkish: Aspects of Turkish-German bilingual children's language use. In Natascha Müller (ed.), *(In)Vulnerable domains in multilingualism* (Hamburg Studies in Multilingualism 1), 221–270. Amsterdam: Benjamins.

Herkenrath, Annette & Jochen Rehbein. 2012. Pragmatic corpus analysis. In Thomas Schmidt & Kai Wörner (eds.), *Multilingual corpora and multilingual corpus analysis* (Hamburg Studies in Multilingualism 14), 123–152. Amsterdam: Benjamins.

Johanson, Lars. 1971. *Aspekt im Türkischen. Vorstudien zu einer Beschreibung des türkeitürkischen Aspektsystems* [Aspect in Turkish: Preliminary studies for a description of

the Turkish aspect system]. (Acta Universitatis Upsaliensis. Studia Turcica Upsaliensia I). Uppsala.
Johanson, Lars. 1973. Sprachbau und Inhaltssyntax am Beispiel des Türkischen [Language construction and content syntax exemplified by Turkish]. *Orientalia Suecana* 22. 82–106.
Johanson, Lars. 1975a. Some remarks on Turkic "Hypotaxis". In Lars Johanson (ed.), *Linguistische Beiträge zur Gesamtturkologie*, 210–224. Budapest: Akadémiai Kiadió.
Johanson, Lars. 1975b. Fiilimsi önermelerin görevleri üzerine [On the functions of non-finite propositions]. In *I. Türk Dili Bilimsel Kurultayına Sunulan Bildiriler* [Proceedings of the First Academic Turkish Linguistic Congress], 525–529. Ankara. Türk Dil Kurumu [Turkish Language Society].
Johanson, Lars. 1990a. Subjektlose Sätze im Türkischen [Subjectless clauses in Turkish]. In Bernt Brendemoen (ed.), *Altaica Osloensia* [Altaic Studies from Oslo], 199–218. Oslo: Universitetsforlaget.
Johanson, Lars. 1990b. Die "problematischen" türkischen Nebensätze [The "problematic" Turkish subordinate clauses]. *Türk Kültürü Araştırmaları* [Turkish Culture Research] 28. 201–209.
Johanson, Lars. 1995. Mehrdeutigkeit in der türkischen Verbalkomposition [Ambiguity in Turkish verbal composition]. In: Marcel Erdal & Semih Tezcan (eds.), *Beläk Bitig: Sprachstudien für Gerhard Doerfer zum 75. Geburstag* [Beläk Bitig: Language Studies for Gerhard Doerfer for his 75th birthday]. (Turcologica 23.), 81–101. Wiesbaden: Harrassowitz.
Johanson, Lars. 1996. Kopierte Satzjunktoren im Türkischen [Copied clause junctors in Turkish]. *Sprachtypologie und Universalienforschung* [Language Typology and Universals] (STUF) 49(1). 39–49.
Johanson, Lars. 2000. Viewpoint operators in European languages. In Östen Dahl (ed.), *Tense and aspect in the languages of Europe*, 27–187. Berlin & New York.
Johanson, Lars. 2006. Two approaches to specificity. In Leonid Kulikov, Andrej Malchukov & Peter de Swart (eds.), *Case, valency and transitivity* (Studies in Language Companion 77), 225–247. Amsterdam: Benjamins.
Johanson, Lars. 2010. Three kinds of clause junctors. In Georg Ziegelmeyer & Norbert Cyffer (eds.), *Aspects of co- and subordination. Case studies from African, Slavonic and Turkic languages*, 9–14. Köln: Rüdiger Köppe.
Johanson, Lars. 2013. Selection of subjunctors in Turkic non-finite complement clauses. *Bilig* 67. 73–86
Johanson, Lars & Éva Á. Csató (eds.). 2006 [1998]. *The Turkic Languages*. London: Routledge.
Karakoç, Birsel. 2005. *Das finite Verbalsystem im Nogaischen* [The finite verbal system in Noghay] (Turcologica 58). Wiesbaden: Harrassowitz.
Karakoç, Birsel. 2007a [2002]. Nogayca ve Türkiye Türkçesinde tümleç yan cümlelerinde yüklemleştiriciler [Predicators in Noghay and Turkish complement clauses]. In László Károly (ed.), *Turcology in Turkey*. Papers selected by Nurettin Demir & Emine Yılmaz, 337–359. Szeged: Altajisztikai Tanszék (Studia uralo-altaica 47).
Karakoç, Birsel. 2007b. Ein Überblick über postverbiale Konverbien im Nogaischen [A survey of postverbial converbs in Noghay]. In Hendrik Boeschoten & Heidi Stein (eds.), *Einheit und Vielfalt in der türkischen Welt. Materialien der 5. Deutschen Turkologenkonferenz, Universität Mainz*, 4. – 7. Oktober 2002 [Unity and diversity in the Turkish world: Materials of the 5th German Conference of Turcologists, University of Mainz, 4–7 October 2002] (Turcologica 69), 215–229. Wiesbaden: Harrassowitz.

Karakoç, Birsel. 2013a. Az konuşurlu bir Türk dili: Nogayca ve Nogaylar üzerine [A Turkic language with few speakers: The Noghay language and the Noghay people]. *Tehlikedeki Diller Dergisi* [Journal of Endangered Languages] 2. 273–301.

Karakoç, Birsel. 2013b. On 'obligatory' control in Turkish and Noghay complement clauses. *Tehlikedeki Diller Dergisi* [Journal of Endangered Languages] 2. 303–323.

Karakoç, Birsel. (forthcoming). Prospectives, proximatives and avertives in Noghay. In: Korn, Agnes & Nevskaya, Irina (eds.) *Prospective and proximative as grammatical categories (Iran–Turan)*. Wiesbaden: Reichert.

Kehayov, Petar & Kasper Boye. This volume. Complementizer semantics in European languages: Overview and generalizations.

Kennelly, Sarah D. 1993. Theta government in Turkish: effects on IP dominated by DP. Ms.

Kennelly, Sarah D. 1994. The semantics of clausal determiners: Subject/ object asymmetry in participles. In G. Borgato (ed.), *Teoria del Linguaggio e Analisi Linguistica* [Theory of Language and Linguistic Analysis], 254–273. Padova: Unipress.

Kerslake, Celia. 2007. Alternative subordination strategies in Turkish. In Jochen Rehbein, Christiane Hohenstein & Lukas Pietsch (eds.), *Connectivity in grammar and discourse* (Hamburg Studies on Multilingualism 5), 231–258. Amsterdam: Benjamins.

Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.

Kornfilt, Jaklin. 2006. Agreement: The (unique and local) syntactic and morphological licenser of subject Case. In João Costa (ed.), *Studies on agreement*, 141–173. Amsterdam: Benjamins.

Kornfilt, Jaklin. 2007. Verbal and nominalized finite clauses in Turkish. In Irina Nikolaeva (ed.), *Finiteness. Theoretical and empirical foundations*, 305–332. Oxford: Oxford University Press.

Kornfilt, Jaklin & John Whitman (eds.). 2012a. *Nominalizations in linguistic theory.* Special issue of *Lingua* 121(7). Amsterdam: Elsevier.

Kornfilt, Jaklin & John Whitman. 2012b. Introduction: Nominalizations in syntactic theory. In Jaklin Kornfilt & John Whitman (eds.), *Nominalizations in linguistic theory. Lingua* 121(7). [Special issue]. 1160–1163.

Kornfilt, Jaklin & John Whitman. 2012c. Afterword: Nominalizations in syntactic theory. In Jaklin Kornfilt & John Whitman (eds.), *Nominalizations in linguistic theory. Lingua* 121(7). [Special issue]. 1297–1313.

Lees, Robert B. 1965. Turkish nominalizations and a problem of ellipsis. *Foundations of Language* 1(2). 112–121.

Lees, Robert B. 1968 [1963]. *The grammar of English nominalizations*, 5th edn. Bloomington and The Hague: Indiana University Press and Mouton.

Noonan, Michael. 2007 [1985]. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 2, *Complex constructions*, 2nd edn., 52–150. Cambridge: Cambridge University Press.

Rehbein, Jochen. 2006. *Ki* – Form and function of a Turkish particle and its contact-induced reinterpretation by bilingual children. Paper presented at the 13th International Conference on Turkish Linguistics, Uppsala University, 16–20 August.

Rehbein, Jochen. 2007. Matrix constructions. In Jochen Rehbein, Christiane Hohenstein & Lukas Pietsch (eds.), *Connectivity in grammar and discourse* (Hamburg Studies on Multilingualism 5), 419–447. Amsterdam: Benjamins.

Rehbein, Jochen, Thomas Schmidt, Bernd Meyer, Franziska Watzke & Annette Herkenrath. 2004. Handbuch für das computergestützte Transkribieren nach HIAT [Handbook for

computer-assisted transcription according to HIAT]. Universität Hamburg, Sonderforschungsbereich 538 Mehrsprachigkeit: Arbeitspapiere zur Mehrsprachigkeit/ Working Papers in Multilingualism, Folge B, Nr. 56.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer.

Rosenbaum, Peter. 1967. *The grammar of English predicate complement constructions*. Cambridge, MA: The MIT Press.

Szabolcsi, Anna. 1992. The noun phrase. Submitted version, available: https://files.nyu.edu/as109/public/NP.pdf (accessed 1 July 2013).

Underhill, Robert. 1976. *Turkish grammar*. (5th printing, 1987). Cambridge, MA: MIT Press.

Van Schaaik, Gerjan. 1999. The order of nominalizations in Turkish. *Turkic Languages* 3. 87–120.

Van Schaaik, Gerjan. 2014. Complications in Turkish complementation: for Éva. In Nurettin Demir, Birsel Karakoç & Astrid Menz (eds.), *Turcology and linguistics. Éva Ágnes Csató festschrift*, 401–416. Ankara: Hacettepe University publications.

Mikhail Knyazev

Complementizers in Kalmyk

1 Introduction

Kalmyk uses as complementizers what appear to be non-finite forms of the verb *gi-* 'say', especially, the converbial form *giǯə* (but also the converbial form *giʁäd* and the participial form *gisən*). According to Noonan (1985), complementizers derived from verbs are rare compared to pronouns, conjunctions, adpositions and case markers. However, they are quite common in the Altaic languages and can also be found in a number of typologically diverse languages listed in Heine & Kuteva (2002).

One interesting fact about Kalmyk complementizers is that there is evidence for treating them *synchronically* as non-finite forms of the verb *gi-* despite the fact that they apparently lack the lexical content of this verb. The other interesting fact is that *contentful* uses of the verb *gi-*, in turn, share a number of syntactic properties with the complementizer *giǯə*. This puzzling combination of facts calls for an analysis.

On a theoretical level the point of this paper is to disentangle verbal and complementizer-like properties of the verb *gi-* and its forms used as complementizers. I will arrive at the somewhat paradoxical conclusion that *gi-* is a verb and a complementizer at the same time, and present a formal implementation of this idea. The major focus of this paper, however, is empirical and it is to document the syntactic and semantic properties and the distribution of the verb *gi-* and its forms used as complementizers.[1]

The structure of this paper is as follows. In this section I give a brief overview of the Kalmyk language (Section 1.1) and its complementation system (Section 1.2). Sections 2 and 3 contain the empirical core of the paper. In Section 2 I describe the general properties of *gi-* as a complement-taking verb. In Section 3 I discuss uses of *gi-* as a complementizer focusing on the form *giǯə*. I describe its general properties (Section 3.1) and its distribution (Section 3.2), including its distributional differences from the complementizer *giʁäd* (Section 3.3). In Section 4 I present my analysis of the verb *gi-* and its forms used as complementizers. Section 4.1 is

1 Kalmyk complementizers have been discussed to some extent in a few descriptive papers (see Baranova 2010, Knyazev 2009, Prokhorov 2009). The discussion there, however, centers around individual aspects of the behavior of the complementizers (such as their distribution, etymological source or interpretation of indexicals). In this paper I provide a more comprehensive description of the data and also present an account of the properties that Kalmyk complementizers have.

a general overview, in sections 4.2 and 4.3 I focus on the syntactic and semantic aspects of the proposed analysis. Section 5 concludes the paper.[2]

1.1 Overview of Kalmyk

In this section I give a brief overview of the Kalmyk language.[3] Kalmyk (Kalmuck) is a Western Mongolic language mainly spoken in the Republic of Kalmykia in the Russian Federation, where it is an official language, alongside Russian. According to a recent census (2010), Kalmyk is spoken by around 80,000 people. Kalmyk is considered an endangered language, according to UNESCO.

Kalmyk has many features in common with other Altaic languages. It is an agglutinative language with vowel harmony. In the nominal domain, Kalmyk lacks gender and a fully grammaticalized number, instead having possessive morphemes, including a possessive reflexive marker and a rich case system with differential object marking. An interesting property of Kalmyk is that it has optional accusative marking of embedded subjects. In the verbal domain, Kalmyk is a subject-verb agreement language with agglutinative tense and aspect markers and a rich set of markers for non-indicative moods. As to the word order, Kalmyk is an SOV language with postpositions. It has a relatively free word order with topicalization and verb phrase internal scrambling as possible options, and lacks obligatory fronting of wh-words.

One feature of Kalmyk particularly important for the subsequent discussion is that it has limited sentential embedding. Most embedded clauses, including both complement and adverbial clauses, are headed by non-finite forms. Yet, sentences, or finite clauses, can be embedded as complements when they are introduced by the forms of the verb *gi-*. It is finite complements that will be the focus of this paper but before I look at them in some detail I will briefly discuss the complementation system of Kalmyk.

---

2 The paper is based on the data obtained in a remote fieldwork setting from a language consultant (born in 1953, a speaker of the Dörbet dialect) in 2011–2013. I am grateful to Olga Seesing for her immense help with the fieldwork. I would also wish to thank the audiences of the Ninth Workshop of Altaic Formal Linguistics (WAFL 9) (Cornell University, 23 August 2013) and the Ling-Lunch at MIT (30 August 2013) for their questions and comments.

3 For a fuller typological overview the reader is referred to Blaesing (2003) and Say et al. (2009)

1.2 Overview of the Kalmyk complementation system

Two major complement types can be identified in Kalmyk based on the morpho-syntactic properties of the predicative head of the complement. The first type is non-finite complements, which are headed by a predicate carrying (a) one of the participial suffixes in their nominalizing function or (b) one of the converbial suffixes.[4] The (a) subtype is illustrated with the complement of verb *üz-* 'see' in (1), where the embedded predicate carries the (contextual allomorph of the) past participle suffix *-sən* and also nominal morphology, namely, case and possessive agreement suffixes. Note that the embedded subject is marked with accusative case.

(1) *bi* [*čamagə* *ger-ür-n* *ir-s-i-čən*]
1SG-NOM 2SG.ACC home-DIR-POSS.REFL come-PTCP.PST-ACC-POSS.2SG
üz-lä-v.
see-REM-1SG
'I saw that you came home.'[5]

The (b) subtype is illustrated with the complement of verb *sed-* 'want' in (2), where the embedded predicate carries the purposive converb suffix *-xar*.

(2) *küükəd* [*nam-ta* *naad-xar*] *sed-əv.*
children 2SG-ASSOC play-CVB.PURP want-PST
'Children wanted to play with me.'

The second type is finite complements, where the verb appears in the indicative or one of the imperative forms.[6] Finite complements also come in two subtypes.

---

4 Since it is not clear how complements of the (b) subtype can be formally distinguished from the corresponding adverbial clauses, their status as a separate complement type remains an open question.

5 I follow the Leipzig Glossing Rules, supplying glosses along the lines of Say et al. (2009) where required. Note that most nouns, except for ones that idiosyncratically require the extension marker *-n* [EXT], have unmarked nominative forms. Since Kalmyk has differential object marking, there are also unmarked accusative forms, which coincide with unmarked nominative forms. However, in line with Say et al. (2009), I do not distinguish the unmarked accusative and the unmarked nominative in the glosses as the two are in complementary distribution, the former being restricted to the object position (hence embedded accusative subjects are always marked by a suffix), whereas the latter is restricted to the subject position.

6 In finite complements the verb may also appear as a participle used finitely. Cf. the finite use of the past participle illustrated in (i).

(i) *küü-n* *ir-sən.*
man-EXT come-PTCP.PST
'The man has come.'
(Blaesing 2003:242)

They are comprised of finite clauses either (a) directly embedded by a matrix verb (the only matrix verb that can directly embed a sentence is *gi-*; see the next section), or (b) introduced by the complementizers such as *giǯə*, *giʁäd* or *gisən*, which are mophologically non-finite forms of the verb *gi-* 'say'. The (a) subtype is illustrated in (3), where a clause headed by a verb in the jussive is embedded by *gi-*.

(3) *eckə-n^{j}* [*Badma* *nan-də* *möngə* *ög-txä*] *gi-və.*
father-POSS.3 Badma 1SG-DAT money give-JUSS say-PST
'His father said Badma should give me the money.'

The (b) subtype is illustrated in (4), where the same embedded clause as in (3) is introduced by the complementizer *giǯə* to become an argument of the verb *kel-* 'tell'.

(4) *eckə-n^{j}* [*Badma* *nan-də* *möngə* *ög-txä* *gi-ǯə*] *kelə-v.*
father-POSS.3 Badma 1SG-DAT money give-JUSS say-CVB.IPFV tell-PST
'His father said that Badma should give me the money.'

In this paper I focus on the forms of *gi-* used as complementizers. I will also discuss uses of *gi-* as a contentful complement-taking verb since these are crucial for understanding the nature and status of the complementizers. An important terminological note is in order. Although I will eventually propose a unified analysis of both uses of *gi-*, for ease of exposition I will refer to forms of *gi-* used as complementizers simply as complementizers and to forms of *gi-* used as a complement-taking verb as simply the verb *gi-*.

2 *Gi-* as a complement-taking verb

In the previous section I noted that *gi-* is able to directly embed finite clauses. In (3) we saw an embedded clause headed by a verb in the jussive. In (5) some other imperative forms are illustrated.

(5) a. *ekə* *eckə* [*namagə* *usə* *av-čə* *irə-Ø*] *gi-və.*
mother father 1SG.ACC water take-CVB.IPFV come-IMP say-PST
'My parents asked me to bring some water.'
(Prokhorov 2009:193)

b. *Badma* [*Ajsa-gə* *Elst-ür* *jov-ij*] *gi-və.*
Badma Ajsa-ACC Elista-DIR go-HORT say-PST
'Badma invited Ajsa to go to Elista.'
(Prokhorov 2009:193)

The embedded verb can also appear in the indicative, as is illustrated in (6a) and (6b).[7] It has to be pointed out that many speakers judge indicative sentences embedded by the verb *gi-* as slightly degraded (and prefer to introduce them with a complementizer).

(6) a. *Ajsa* [*udlgo* *aavə-m* *ir-xə*] *gi-və.*
Ajsa soon grandfather-POSS.1SG come-PTCP.FUT say-PST
'Ajsa said my grandpa will come soon.'
b. *Ajsa* [*ekə-m* *kövü* *ʁarʁ-la*] *gi-və.*
Ajsa mother-POSS.1SG son give_birth-REM say-PST
'Ajsa said my mother has had a boy.'

The embedding of finite clauses under the verb *gi-* has the following properties. Firstly, *gi-* is the only verb that is able to embed a sentence, as is demonstrated in (7) by the unacceptability of replacing the verb with *kel-* 'tell'. To make (7) acceptable, one should plug in a complementizer such as *giǯə*, as in (4). Similar facts are replicated for indicative embedded clauses, as is shown in (8a) and (8b).

(7) **eckə-n^{j}* [*Badma* *nan-də* *möngə* *ög-txä*] *kelə-v.*
father-POSS.3 Badma 1SG-DAT money give-JUSS tell-PST
intended: 'His father said Badma should give me the money.'

(8) a. **Ajsa* [*udlgo* *aavə-m* *ir-xə*] *kelə-v.*
Ajsa soon grandfather-POSS.1SG come-PTCP.FUT tell-PST
intended: 'Ajsa said my grandpa will come soon.'
b. *Ajsa* [*udlgo* *aavə-m* *ir-xə* *gi-ǯə*] *kelə-v.*
Ajsa soon grandfather-POSS.1SG come-PTCP.FUT say-CVB.IPFV tell-PST
'Ajsa said my grandpa will come soon.'

Secondly, embedded clauses introduced by *gi-* must immediately precede the verb. Thus neither a dative nor a subject DP may intervene between the clause and *gi-*, as can be seen in (9a) and (10a). Note that this is possible with complements introduced by the complementizer, as can be seen in (9b) and (10b); cf. also (10c), which shows that a dative DP is in principle compatible with a sentence embedded by the verb *gi-*.[8]

(9) a. *[*aavə-m* *udlgo* *ir-xə*] *Ajsa* *gi-və.*
grandfather-POSS.1SG soon come-PTCP.FUT Ajsa say-PST
intended: 'Ajsa said my grandpa will come soon.'

7 Note that the future participle in (6a) is used finitely. See footnote 6.

8 A question arises whether the immediate precedence is a property of the construction with an embedded sentence or of the verb *gi-* in general, in which case we would expect that a nominal complement of *gi-* would also immediately precede it. This is somewhat hard to check since, as I will show shortly, speakers usually reject examples with *gi-* taking nominal complements.

b. [*aavə-m* *udlgo* *ir-xə* *gi-ǯə*] *Ajsa*
grandfather-POSS.1SG soon come-PTCP.FUT say-CVB.IPFV Ajsa tell-PST
'Ajsa said that my grandpa will come soon.'

(10) a. **Ajsa* [*aavə-m* *ir-xə*] *Badma-də* *gi-və.*
Ajsa grandfather-POSS.1SG come-PTCP.FUT Badma-DAT say-PST
intended: 'Ajsa said to Badma that my grandpa is coming.'
b. *Ajsa* [*aavə-m* *ir-xə* *gi-ǯə*] *Badma-də*
Ajsa grandfather-POSS.1sg come-PTCP.FUT say-CVB.IPFV Badma-DAT
kelə-v.
tell-PST
'Ajsa said to Badma that my grandpa is coming.'
c. *Ajsa* *Badma-də* [*aavə-m* *ir-xə*] *gi-və.*
Ajsa Badma-DAT grandfather-POSS.1SG come-PTCP.FUT say-PST
'Ajsa said to Badma that my grandpa is coming.'

Thirdly, *gi-* disallows nominal complements. This is illustrated for the wh-word *jun* 'what' in the unmarked form, as in (11a); cf. the corresponding example with the verb *kel-* 'tell', as shown in (11b). The same holds for participial forms used as nominalizations, as shown in (12a); cf. (12b).

(11) a. **Bajrta* *čam-də* *ju* *gi-və?*
Bajrta 2SG-DAT what say-PST
intended: 'What did Bajrta say to you?'
b. *Bajrta* *čam-də* *ju* *kelə-v?*
Bajrta 2SG-DAT what say-PST
'What did Bajrta say to you?'

(12) a. **Bajrta* [*bi* *xotə* *säänär* *ke-dg-igə*] *gi-və.*
Bajrta 1SG.NOM food well do-PTCP.HAB-ACC say-PST
intended: 'Bajrta said that I cook well.'
b. *Bajrta* [*bi* *xotə* *säänär* *ke-dg-igə*] *kelə-v.*
Bajrta 1SG.NOM food well do-PTCP.HAB-ACC say-PST
'Bajrta said that I cook well.'

The fourth property of *gi-* is related to the interpretation of the complement clause. A clause introduced by *gi-* allows indexicals inside it to be interpreted both relative to the context of utterance and to the reported context. For example, the first person pronoun *nandə* in sentence (13), repeated from above, could refer either to the speaker, as in (13i), or to the subject, as in (13ii).[9] The same is true of the first person possessive marker *-m* in example (14) with the indicative.

9 It is worth mentioning that the the interpretation of indexicals relative to the reported context does not necessarily involve direct speech or quotation. This can be shown by the fact that this intepretation can obtain even when the subject of the complement clause is marked with ac-

(13) *eckə-nʲ* [*Badma* *nan-də* *möngə* *ög-txä*] *gi-və*.
father-POSS.3 Badma 1SG-DAT money give-JUSS say-PST
i. 'His father said Badma should give me the money.'
ii. 'His father$_i$ said Badma should give him$_i$ the money.'

(14) *Ajsa* [*udlgo* *aavə-m* *ir-xə*] *gi-və*.
Ajsa soon grandfather-POSS.1SG come-PTCP.FUT say-PST
i. 'Ajsa said that my grandpa will come soon.'
ii. 'Ajsa$_i$ said that her$_i$ grandpa will come soon.'

To summarize, *gi-* is the only verb in Kalmyk that can embed clausal complements directly, it has to immediately precede the complement clause, it disallows nominal complements and it allows indexicals inside the complement clause to be interpreted both relative to the context of utterance and to the reported context.

Before concluding this section, it is important to note that *gi-* can also be used outside complementation both as a functional element and as a contentful verb. As a functional element, (converbial forms of) *gi-* can introduce purpose clauses, as shown in (15) for *giʁäd*.[10] Secondly, *gi-* can also be used in a construction with the modal converb expressing the avertive meaning, as shown in (16). Finally, non-finite forms of *gi-* can introduce ideophonic elements and proper names, as shown in (17a) and (17b), respectively. As a contentful verb, *gi-* can introduce onomatopoeic words, as shown in (18).

(15) [*čamagə* *tövknün* *gerg-tä-ʁän* *küünd-txä* *gi-ʁad*]
2SG.ACC comfortable wife-ASSOC-POSS.REFL speak-JUSS say-CVB.PFV
bi *jov-x-u*.
1sg.nom go-PTCP.FUT-1SG
'I will go so that you could comfortably speak to your wife.'
(Say 2009: 433)

(16) *Badma* *unə-n* *gi-və*.
Badma fall-CVB.MOD say-PST
'Badma was on the verge of falling.'
(Baranova 2010)

---

cusative case, indicating that the clause is embedded. This is shown in (i). This phenomenon is referred to in the literature as *indexical shift*; see Shklovsky & Sudo (2013).

(i) *eckə-nʲ* [*Badma-gə* *nan-də* *möngə* *ög-txä*] *gi-və*.
father-POSS.3 Badma-ACC 1SG-DAT money give-JUSS say-PST
a. 'His father said Badma should give me the money.'
b. 'His$_i$ father said Badma should give him$_i$ the money.'

10 Heine & Kuteva (2002) cite this as a possible path of grammaticalization of 'say'.

(17) a. *gils-gils* *gi-ǯə* *xälä-*
IDEO say-CVB.IPFV look
'look with eyes cast down'
(Baranova 2010)

b. *eckə-m* *nan-də* *Ivan* *gi-dəg* *nerə* *ogə-v.*
father-POSS.1SG 1SG-DAT Ivan say-PTCP.HAB name give-PST
'Father gave me Ivan as a name.'
(Baranova 2010)

(18) *ʁav-ʁav* *gi-*
bow-wow say
'bark'
(Baranova 2010)

I will not discuss these uses of *gi-* any further and will ignore them in the proposed account of *gi-*, limiting myself to contexts where *gi-* introduces complement clauses – either as a contentful verb, as we just saw in this section, or as a complementizer, as we will see in the following section.

3 *Gi-* as a complementizer

As I mentioned above, Kalmyk uses non-finite forms of the verb *gi-* 'say' as complementizers, such as *giǯə*, *giʁäd* and *gisən*. More precisely, complementizer *giǯə* is derived by the imperfective converb suffix *-ǯə*, complementizer *giʁäd* is derived by the perfective converb suffix *-ad*, complementizer *gisən* is derived by the past participle suffix *-sən* (see Say et al. 2009) for the meaning of the relevant suffixes). In this paper I will focus on the complementizer *giǯə*, which is illustrated in (19), repeated from above, with the embedded verb in the jussive; for examples with the embedded verb in the indicative and the imperative see, e.g., (8b) and (26), respectively.

(19) *eckə-n*^{j} [*Badma* *nan-də* *möngə* *ög-txä* *gi-ǯə*] *kelə-v.*
father-POSS.3 Badma 1SG-DAT money give-JUSS say-CVB.IPFV tell-PST
'His father said that Badma should give me the money.'

I will limit my discussion of *giʁäd* to its distributional differences with *giǯə* and only briefly touch on *gisən*.

3.1 General properties

In this section I will discuss general morphosyntactic and semantic properties of the complementizer *giǯə*. Firstly, complementizers morphologically related to the verb *gi-*, such as *giǯə*, *giʁäd*, *gisən* and some other derivates of *gi-* (see footnote 12) are the only ones that can introduce finite complement clauses.

Secondly, the complementizer *giǯə* has to immediately precede the complement clause that it introduces, just like the verb *gi-*. Thus, for example, a dative DP cannot intervene between the clause and the complementizer, as can be seen in (20). This adjacency requirement might appear a self-evident property of complementizers. Yet, given that Kalmyk complementizers are morphologically related to a verb, one might in principle expect that they would share some behavior with ordinary verbs such as *kel-* 'tell', which allow dative DPs to intervene between them and their complements, as we saw in (10b).

(20) *[*Ajsa* *aavə-m* *ir-xə*] *Badma-də* *gi-ǯə* *kelə-v.*
Ajsa grandfather-POSS.1SG come-PTCP.FUT Badma-DAT say-CVB.IPFV tell-PST
intended: 'Ajsa said to Badma that my grandpa is coming.'

Thirdly, complement clauses introduced by *giǯə* allow indexicals to be interpreted both relative to the context of utterance and to the reported context, as shown in (21) and (22).[11]

(21) *eckə-n*^{j} [*Badma* *nan-də* *möngə* *ög-txä* *gi-ǯə*] *kelə-v.*
father-POSS.3 Badma 1SG-DAT money give-JUSS say-CVB.IPFV tell-PST
i. 'His father said that Badma should give me the money.'
ii. 'His father$_i$ said that Badma should give him$_i$ the money.'

(22) *Ajsa* [*udlgo* *aavə-m* *ir-xə* *gi-ǯə*] *kelə-v.*
Ajsa soon grandfather-POSS.1SG come-PTCP.FUT say-CVB.IPFV tell-PST
i. 'Ajsa said my grandpa will come soon.'
ii. 'Ajsa$_i$ said her$_i$ grandpa will come soon.'

The fourth property of the complementizer *giǯə* is that it cannot introduce complements of nouns. This is illustrated with *zakvər* 'order' and *zäŋgə* 'news' in (23a)

11 Cf. (i) indicating that complement clauses introduced by *giǯə* also allow indexical shift; see footnote 9.
(i) *eckə-n*^{j} [*Badma-gə* *nan-də* *möngə* *ög-txä* *gi-ǯə*] *kelə-v.*
father-POSS.3 Badma-ACC 1SG-DAT money give-JUSS say-CVB.IPFV tell-PST
a. 'His father said Badma should give me the money.'
b. 'His$_i$ father said Badma should give him$_i$ the money.'

and (23b). Instead of *gižə* the form *gisən* is used, which is morphologically the past participle of *gi-*.[12] This is shown in (24a) and (23b).

(23) a. **cergč-nər* [*ʁolə* *tal* *jov-tən* *gi-žə*] *zakvər* *avə-v.*
soldier-PL river towards go-IMP.PL say-CVB.IPFV order receive-PST
intended: 'Soldiers received the order to go towards the river.'
b. *[*tüümər* *šatə-v* *gi-žə*] *zäŋgə* *terə* *soŋsə-v.*
fire burn-PST say- CVB.IPFV news that hear-PST
intended: 'He heard the news that a fire broke out.'

(24) a. [*tüümər* *šatə-v* *gi-sən*] *zäŋgə* *terə* *soŋsə-v.*
fire burn-PST say-PTCP.PST news that hear-PST
'He heard the news that a fire broke out.'
b. *cergč-nər* [*ʁolə* *tal* *jov-tən* *gi-sən*] *zakvər* *avə-v.*
soldier-PL river towards go-IMP.PL say-PTCP.PST order receive-PST
'Soldiers received the order to go towards the river.'

Assuming that converbs only attach to a verbal predication, the incompatibility of *gižə* and complements of nouns suggests that *gižə* displays distributional restrictions typical of converbs, whereas *gisən* displays distributional restrictions typical of participles.

To summarize, complementizers related to the verb *gi-* such as *gižə* essentially share syntactic properties with the verb *gi-* (in their unique ability to introduce complement clauses, in the adjacency requirement and the interpretation of indexicals). This would suggest that the verb *gi-* and the complementizers should receive a unified analysis. This conclusion is reinforced by the fact that the complementizers syntactically pattern as the respective non-finite forms, which justifies their treatment as non-finite forms of the verb *gi-* synchronically. However, the complementizer *gižə* dramatically differs from the verb *gi-* in that it lacks lexical content as manifested by its distribution across complement-taking predicates, to which I now turn.

12 Other derivates that can introduce clausal complements of nouns are *gižäsən* and *gižäxə*, illustrated in (ia) and (ib).

(i) a. *Bajrta-də* [*kövü-n^{j}* *institut-tə* *sur-xə* *gi-žä-sən*]
Bajrta-DAT son-POSS.3 institute-DAT study-PTCP.FUT say-DUR-PTCP.PST
itkəl *bi-lä.*
belief be-rem
'Bajrta had the belief that her son would study at the institute.'
b. [*tüümər* *šatə-v* *gi-žä-xə*] *zäŋgə* *irə-v.*
fire burn-PST say-DUR-PTCP.FUT news come-PST
'The news came that a fire broke out.'

3.2 Distribution across complement-taking predicates

So far we have seen examples where *gižə* introduces sentential complements of the verb *kel-* 'tell', which takes both indicative and imperative complements. Other utterance predicates[13] include *zäŋgl-* 'announce', which takes indicative complements, and *sur-* 'ask' and *zak-* 'order', which take imperative complements. These verbs are illustrated in (25), (26) and (27). Note that imperative complements in (26) and (27) are rendered with infinitive in English, which suggests that the verbs in those cases can also function as *manipulative* predicates (see Noonan 1985).

(25) *axlač* [*tüümər* *šatə-v* *gi-žə*] *zäŋglə-v.*
chairman fire burn-PST say-CVB.IPFV announce-PST
'The chairman announced that a fire broke out.'

(26) *Badma* *kövü-ʁän* [*du* *duul-Ø*] *gi-žə* *surə-v.*
Badma son-POSS.REFL song sing-IMP say-CVB.IPFV ask-PST
'Badma asked his son to sing a song.'

(27) *Baatər* *cergč-nər-tə* [*ʁolə* *tal* *jov-tən*] *gi-žə* *zakə-v.*
Baatr soldier-PL-DAT river towards go-IMP.PL say-CVB.IPFV order-PST
'Baatr ordered soldiers to go towards the river.'

Importantly, examples with utterance predicates do not allow us to determine whether *gižə* has an independent semantic contribution. This is because speech acts, which are encoded by utterance predicates, in general have multiple semantic parameters such as the content, manner, goal, etc. of the speech act, and it is difficult to say in advance what parameters a given verb lexicalizes. As a result, it might be that various parameters are not exclusively lexicalized by the utterance predicate (in which case *gižə* would be lacking semantic contribution) but rather distributed across that predicate and *gižə* in a given construction. For example, this would be the case if *gižə* lexicalized the semantic feature corresponding to the transfer of information (the essential component of utterance predicates) whereas the predicate lexicalized the rest of the features associated with the speech act.

The distribution of *gižə*, however, is not limited to utterance predicates and extends to predicates that lack the 'transfer of information' component to their meaning. Thus, for example, *gižə* can introduce complements of *propositional attitude* predicates such as *san-* 'think' and *almac-* 'doubt', illustrated in (28) and (29), and the verb *ää-* 'fear', illustrated in (30).

13 Hereafter I use the terminology of Noonan (1985) to refer to the classes of complement-taking predicates.

(28) *bi* [*Ajsa-gə* *degtər* *umš-xə*] *gi-ǯə* *san-ǯa-na-v.*
1SG.NOM Ajsa-ACC book read-PTCP.FUT say-CVB.IPFV think-DUR-PRS-1SG
'I think Ajsa will read a book.'

(29) [*enə* *naaran* *ir-xə*] *gi-ǯə* *bi* *almac-ǯa-na-v.*
this here come-PTCP.FUT say-CVB.IPFV 1SG.NOM doubt-DUR-PRS-1SG
'I doubt that he will come here.'

(30) *Bajərta* [*kövü-ʁän* *cerg-tə* *mord-xə*] *gi-ǯə* *ää-ǯä-nä.*
Bajrta son-POSS.REFL army-DAT leave-PTCP.FUT say-CVB.IPFV fear-DUR-PRS
'Bajrta fears that her son will go to the army.'[14]

Giǯə is also compatible with the verb *nääl-* 'hope', which is classified as a *desiderative* predicate in Noonan (1985). The example is shown in (31). Interestingly, the more common desiderative verb *sed-* 'want' resists complements introduced by *giǯə* and instead takes non-finite complements headed by the *-xar* converb, as shown in (32a) and (32b); see also (6) above.[15]

(31) [*kövü-ʁän* *institut-tə* *sur-xə*] *gi-ǯə* *Bajərta*
son-POSS.REFL institute-DAT study-PTCP.FUT say-CVB.IPFV Bajrta
nääl-ǯä-nä.
hope-DUR-PRS
'Bajrta hoped that her son would study at the institute.'

(32) a. **bi* [*čamagə* *zurg* *zur-txa*] *gi-ǯə* *sed-ǯä-nä-v.*
1SG.NOM 2SG.ACC picture draw-JUSS say-CVB.IPFV want-DUR-PRS-1SG
intended: 'I want you to draw a picture.'
b. *bi* [*čam-ar* *zurg* *zur-ul-xar*] *sed-ǯä-nä-v.*
1SG.NOM 2SG-INS picture draw-CAUS-CVB.PURP want-DUR-PRS-1SG
'I want you to draw a picture.'

Another verb *giǯə* can combine with is *soŋs-* 'hear', illustrated in (33). As suggested by the use of a *that*-clause in the English translation, *soŋs-* 'hear' is used as an (*acquisition of*) *knowledge* predicate rather than an immediate perception predicate (which would require a bare infinitive complement). Interestingly, the other perception verb *üz-* 'see' disallows complements introduced by *giǯə* even though the verb is compatible with the acquisition of knowledge use.[16] Instead,

14 Note that the possessive reflexive marker *-an* glossed as 'POSS.REFL' is restricted to accusative contexts when not accompanied by a separate case suffix.

15 Note that in (32b) the converb has a causative suffix, which indicates that the action encoded by the embedded verb is performed by a participant distinct from the matrix subject. See Say (2009) for more on this construction.

16 This is shown by the fact that *üz-* 'see' can take complements with the negated embedded verb, as illustrated in (i). According to Mittwoch (1990), negation is incompatible with immediate

üz- 'see' takes a non-finite complement headed by a participial form. This is shown in (34a) and (34b).

(33) *Badma* [*selän-də* *tüümər* *šatə-v*] *gi-ǯə* *soŋs-la.*
Badma village-DAT fire burn-PST say-CVB.IPFV hear-PST
'Badma heard that a fire broke out in the village.'

(34) a. **bi* [*čamagə* *ger-ürə-n* *irə-v*] *gi-ǯə*] *üz-lä-v.*
1SG.NOM 2SG.ACC home-DIR-POSS.REFL come-PST say-CVB.IPFV see-REM-1SG
intended: 'I saw that you came home.'
b. *bi* [*čamagə* *ger-ürə-n* *ir-s-i-čən*]
1SG.NOM 2SG.ACC home-DIR-POSS.REFL come-PTCP.PST-ACC-POSS.2SG
üz-lä-v.
see-REM-1SG
'I saw that you came home.'

Other knowledge predicates also resist taking sentential complements introduced by *giǯə*. This is, for example, the case with *mart-* 'forget', which instead takes a non-finite complement, as shown in (35a) and (35b). The verb *med-* 'know' also usually takes a non-finite complement, given in (36a). When confronted with sentences involving a complement introduced by *giǯə*, speakers either reject them or report a shift in meaning and note that the verb should then be translated as 'think', as shown in (36b).

(35) a. **terə* [*čamagə* *mör-i-nʲ* *edgä-və*] *gi-ǯə* *mart-čkə-v.*
that 2SG.ACC horse-ACC-POSS.3 cure-PST say-CVB.IPFV forget-COMPL-PST
intended: 'He forgot that you cured his horse.'
b. *terə* [*čamagə* *mör-i-nʲ* *edgä-s-igə*] *mart-čkə-v.*
that 2SG.ACC horse-ACC-POSS.3 cure-PTCP.PST-ACC forget-COMPL-PST
'He forgot that you cured his horse.'

(36) a. [*bi* *enü-gə* *zalu-ʁas-nʲ* *sal-s-i-nʲ*] *med-lä-v.*
1SG.NOM this-ACC husband-ABL-POSS.3 split-PTCP.PST-ACC-POSS.3 know-REM-1SG
'I know that she has split from her husband.'
b. *ʔbi* [*enü-gə* *zalu-ʁasə-nʲ* *sal-ǯə* *od-sən*] *gi-ǯə*
I this-ACC husband-ABL-POSS.3 split-CVB.IPFV go-PTCP.PST say-CVB.IPFV
med-lä-v.
know-REM-1SG
'I think that she has split from her husband.'

---

perception complements. Hence, the example should involve the acquisition of knowledge use of the perception verb.

(i) [*xora-də* *šamə* *esə* *untər-ǯa-x-i-nʲ*] *bi* *üz-lä-v.*
room-DAT light NEG fade-DUR-PTCP.FUT-ACC-POSS.3 1SG.NOM see-REM-1SG
'I saw that the light was not fading.'

When we look at the pattern displayed by knowledge predicates, we observe that 'hear' and 'think', compatible with *gižə*, do not presuppose the truth of their complement as opposed to 'see' and 'know' (and also 'forget'), incompatible with *gižə*, which do presuppose the truth of their complement, that is, are *factive*, as dubbed by Kiparsky & Kiparsky (1970).[17] Thus we can expect that *gižə* would generally resist factive complements.

Consider in this light the data on *commentative* predicates, which are often classified as (emotive) factive predicates and hence expected to be incompatible with *gižə*. Interestingly, this is not the case, as demonstrated by examples in (37a) and (38a), which involve the verbs *bajrəl-* 'be glad' and *ʁund-* 'be offended' functioning as commentative predicates. Note that these verbs also take non-finite complements, as shown in (37b) and (38b).

(37) a. *enə* [*Badma-gə* *bičəg* *bič-žə*] *gi-žə* *bajrəl-žə-na.*
this Badma-ACC letter write-EVID say-CVB.IPFV rejoice-DUR-PRS
'He is glad that Badma wrote a letter.'

b. *enə* [*Badma-gə* *bičəg* *bič-sən-də*] *bajrəl-žə-na.*
this Badma-ACC letter write-PTCP.PST-DAT rejoice-DUR-PRS
'He is glad about the fact that Badma wrote a letter.'

(38) a. [*terə* *institutə* *tögsä-sən* *uga*] *gi-žə* *ʁund-žə-na-v.*
that institute finish-PTCP.PST NEG.COP say-CVB.IPFV be_offended-DUR-REM-1SG
'I regret that he has not finished the university.'

b. [*terə* *institutə* *esə* *tögsä-sən-də-nʲ*] *ʁund-žə-na-v.*
that institute NEG finish-PTCP.PST-DAT-POSS.3 be_offended-DUR-REM-1SG
'I regret the fact that he has not finished the university.'

However, when we look more closely at the difference in meaning between examples (a) and (b) in (37) and (38), we note that the verbs in (37a) and (38a) with *gižə* should in fact be more appropriately translated as 'glad/regret to believe'. This is indicated by the felicity of the continuations in (39a) and (39b) for (37a) and (38b). These continuations ensure that the truth of the complement is not presupposed and thus require a non-factive interpretation of the sentence. Note that for the corresponding examples with participial complements in (37b) and (38b) the relevant continuations are not felicitous.

(39) a. *...bolvə* *terə* *bič-sən* *uga* *bi-lä.*
but that write-PTCP.PST NEG.COP be-REM
'...but (it turned out that) he hadn't write it.'

17 For the observation about 'see' (with a *that*-clause complement) see Higginbotham (1983). 'Know' is usually analyzed as semi-factive as it loses its factivity in certain contexts such as questions and conditionals; see Hooper & Thompson (1973) and references therein. I am grateful to the reviewer for drawing my attention to this fact.

b. *...bolvə terə tögsä-lä.*
but that finish-REM
'...but (it turned out that) he had finished it.'

The contrast between the interpretation of complements introduced by *giǯə* and that of non-finite complements in the case of commentative predicates is very similar to what we saw in the case of *med-* 'know', which reinforces the conclusion that complements introduced by *giǯə* have to be non-factive.

It is important to note that although *giǯə* is incompatible with factive verbs, it can be used to introduce propositions that are taken to be true by virtue of the immediate context, that is, contextually given, as shown in (40). I conclude then that the non-factivity requirement induced by *giǯə* invokes the notion of factivity as a property of the complement-taking verb (or its particular sense) rather than the property of the concrete sentence.[18]

(40) *Badma* [*tüümər* *šatə-v*] *gi-ǯə* *kel-sən* *uga.*
Badma fire burn-PST say-CVB.IPFV say-PTCP.PST NEG.COP
'{Even though he knew it...} Badma did't say that a fire broke out.'

To summarize, *giǯə* is compatible with a wide range of complement-taking predicates including those that do not express speech acts, which suggests that it lacks the 'say' interpretation and has an impoverished semantic content. In addition, *giǯə* displays the non-factive restriction. The non-factive restriction on *giǯə* follows naturally from the fact that *giǯə* is derived from the verb *gi-* 'say', whereas 'say' is classified as non-factive, as in Hooper & Thompson (1973).[19]

We may wonder whether *giǯə* inherits the other semantic feature of the verb 'say', namely, *assertivity*. Following Sheehan & Hinzen (2011:33), I understand assertive complements as "taken to be true by the grammatical subject, and perhaps understood by the hearer as if asserted by the matrix subject". On this understanding assertivity and non-factivity are two independent features as suggested by the existence of assertive verbs that are factive such as 'know' (but see footnote 17) and non-factive verbs that are non-assertive such as 'doubt', 'order' and 'ask'. In this section we saw that *giǯə* is compatible with *sur-* 'ask' (26), *zak-* 'order' (27) and *almac-* 'doubt' (29), which are non-assertive. Thus we can conclude that *giǯə* does not inherit assertivity from *gi-* 'say' and, more generally, is

18 This understanding of factivity is similar to the one used by Haegeman & Ürögdi (2010), who clearly distinguish factivity as a lexico-semantic property of verbs and contextually defined notions such as givenness. I am grateful to the reviewer for bringing up this point.
19 In Section 4.3 this idea is slightly modified.

orthogonal to the assertive/non-assertive distinction. Interestingly, this might not be the case with the other complementizer *giʁäd*, to which I now turn.

3.3 Distributional differences between *giǯə* and *giʁäd*

Except for complements of nouns, which require participial forms such as *gisən*, verbal complements are introduced by converbial forms of the verb *gi-*. So far we have seen examples with *giǯə*, which is derived by the imperfective converb suffix. The other form is *giʁäd*, which is derived by the perfective converb suffix. The distribution of *giʁäd* is very close to that of *giǯə*, but appears to be more restricted.

All speakers accept examples where *giʁäd* introduces complements of the verb *kel-* 'tell' with a non-indicative complement, shown in (41), manipulative verbs like *sur-* 'ask', *zak-* 'order', shown in (42) and (43), and verbs *ää-* 'fear' and *nääl-* 'hope', shown in (44) and (45). Cf. examples with *giǯə* and the corresponding verbs in (4), (26), (27), (30) and (31).

(41) *eckə-n*^{j} [*Badma nan-də möngə ög-txä*] *gi-ʁäd* *kelə-v.*
father-POSS.3 Badma 1SG-DAT money give-JUSS say-CVB.PFV tell-PST
'His father said that Badma should give me the money.'

(42) *Badma* [*kövü-ʁän* *du* *duul-Ø*] *gi-ʁäd* *surə-v.*
Badma son-POSS.REFL song sing-IMP say-CVB.PFV tell-PST
'Badma requested that his son sing a song.'

(43) *Baatr* *cergč-nər-tə* [*ʁolə* *tal* *jov-tən*] *gi-ʁäd* *zakə-v.*
Baatr soldier-PL-DAT river towards go-IMP.PL say-CVB.PFV order-PST
'Baatr ordered soldiers to go towards the river.'

(44) *Bajrta* [*kövü-ʁän* *cerg-tə* *mord-xə*] *gi-ʁäd* *ää-ǯä-nä.*
Bajrta son-POSS.REFL army-DAT leave-PTCP.FUT say-CVB.PFV fear-DUR-PRS
'Bajrta fears that her son will go to the army.'

(45) [*čamagə* *edgä-txä*] *gi-ʁäd* *bi* *nääl-ǯä-nä-v.*
2SG.ACC get_cure-JUSS say-CVB.PFV 1SG.NOM hope-DUR-PRS-1SG
'I hope that you will get cured.'

With other verbs, however, speakers generally reject *giʁäd* and instead prefer *giǯə*. This is illustrated for *kel-* 'tell' with an indicative complement in (46), for *zäŋgl-* 'announce' in (47), for *san-* 'think' in (48) and for *med-* in the sense 'think' in (49). Cf. the corresponding examples with *giǯə* in (8b), (23b), (28) and (36b). The acceptability judgments are quite subtle but there is a perceptible contrast between *giʁäd* and *giǯə* in these cases.

(46) *??Ajsa* [*udlgo* *aavə-m* *ir-xə*] *gi-ʁäd* *kelə-v.*
Ajsa soon grandfather-POSS.1SG come-PTCP.FUT say-CVB.PFV tell-PST
intended: 'Ajsa said my grandpa will come soon.'[20]

(47) *??axlač* [*tüümər* *šatə-v*] *gi-ʁäd* *zäŋglə-v.*
chairman fire burn-PST say-CVB.PFV announce-PST
intended: 'The chairman announced that a fire broke out.'

(48) *?*bi* [*Ajsa-gə* *degtər* *umš-xə*] *gi-ʁäd* *san-ǯa-na-v.*
1SG.NOM Ajsa-ACC book read-PTCP.FUT say-CVB.PFV think-DUR-PRS-1SG
intended: 'I think Ajsa will read a book.'

(49) **bi* [*enü-g* *zalu-ʁasə-nʲ* *sal-ǯə* *od-sən*]
1SG.NOM this-ACC husband-ABL-POSS.3 split-CVB.IPFV go-PTCP.PST
gi-ʁäd *med-lä-v.*
say-CVB.IPFV know-REM-1SG
intended: 'I think that she has split from her husband.'

The question is what is,what makes the verbs in (46)–(49) resist taking *giʁäd*. All verbs in these examples introduce assertions, as opposed to verbs in (41), (42), (43), (44) and (45), which express notions generally associated with non-assertive contexts; see Noonan (1985). This is definitely true for commands and requests in (41), (42) and (43) but can also be extended to predicates of fearing and hope in (44) and (45), as those have a clear emotional component (desire or lack thereof) to their meaning.[21]

In view of these facts I would like to hypothesize that *giʁäd* is restricted to non-assertive contexts. If this hypothesis is on the right track, then in Kalmyk the assertive vs. non-assertive distinction is (optionally) marked by the choice of the complementizer, as opposed to languages where this distinction is marked by the choice of the indicative or non-indicative mood; see Noonan (1985).

It is certainly the case that the choice of mood is related to whether the context is assertive or non-assertive. We saw that with the verb *kel-* 'tell', which can render the context non-assertive once used with an non-indicative mood (by hypothesis).[22] However, the choice of mood is in principle orthogonal to the

20 Speakers sometimes report that it is easier to get the interpretation where the first person possessive marker in *aavəm* refers to the matrix subject. I suspect that this is due to the fact that under this interpretation the complement clause can be construed as a quotation.

21 According to Noonan (1985), predicates of fearing can be doubly classified on the assertive/non-assertive scale depending on their degree of certainty but the important thing is that they are in principle compatible with non-assertive contexts.

22 The reviewer raises an objection according to which the semantic distinction between indicative vs. non-indicative uses of the verb *kel-* 'tell' in (41) and (46) should be analyzed in terms of mood rather than assertion. In my view, however, these two possibilities do not exclude each

assertive vs. non-assertive distinction. This is because the verb *ää-* 'fear', which takes an indicative complement, can take a non-assertive complement (again by hypothesis).

At this point I don't have a clear understanding of why *giʁäd* is restricted to non-assertive contexts, whereas *giǯə* is not (both complementizers being non-factive). Under the account that I will propose below, according to which complementizers should be viewed as non-finite forms of the verb *gi-*, the distributional difference between *giʁäd* and *giǯə* should follow from the difference in the choice of the converbial suffix. Yet there is no clear connection between assertivity vs. non-assertivity and perfectivity vs. imperfectivity (of the converb). I leave this puzzle to future research.

4 A unified analysis

4.1 General considerations

In sections 2 and 3 I looked at the properties of the verb *gi-* and its forms used as complementizers. Taken together they seem to lead us to an apparent paradox. On the one hand, the shared syntactic properties of the verb *gi-* and the complementizers and the fact that the complementizers syntactically pattern with non-finite forms suggest that the complementizers should be treated as forms of the verb *gi-*. On the other hand, the fact that the complementizers could not have the 'say' interpretation as opposed to the verb *gi-*, suggests that, from the point of view of semantics, the verb *gi-* and the complementizers should be treated separately.

To resolve this paradox, I will pursue the idea that the complementizers are in fact forms of the verb *gi-* and explore ways to account for the semantic difference between the two. Two theoretical options present themselves, which I will consider in turn. Firstly, the root *gi-* used as a verb and the root *gi-* used in the complementizers could be analyzed as diachronically related but distinct lexical items with distinct semantic content – full lexical content in the former case and impoverished lexical content in the latter case. Alternatively, *gi-* in both syntactic configurations could be analyzed as a single lexical item.

---

other if assertion is viewed as an abstract semantic feature associated with certain contexts of clausal complementation rather than a purely pragmatic notion having to do with what is presupposed in the sentence.

The first analysis could be couched in terms of *grammaticalization*. As a gradual process of loss of lexical properties (in particular, semantic erosion), grammaticalization could explain why the verb *gi-* lost its lexical content to become a complementizer. In addition, a grammaticalization account could explain why complementizers retained some of the semantic features associated with the semantic content of the source verb such as non-factivity. Moreover, as a multi-faceted process (see Lehmann 1995) grammaticalization does not have to target various morphosyntactic features of the source element all at once, so we can in principle have an item undergoing the process of grammaticalization that lost its lexical content without losing its categorial information and the syntactic distribution associated with it. A grammaticalization account has in fact been proposed for Kalmyk complementizers (see Baranova 2010). For grammaticalization accounts of other complementizers derived from the verb 'say' see Heine & Kuteva (2002).

Despite its attractiveness, a grammaticalization account has one conceptual problem. Under this account, the relevant data taken from a purely synchronic perspective require the postulation of two homophonous morphemes *gi-* related by a grammaticalization process. One *gi-* would be a verb with lexical content, while the other would be a bound root with an impoverished semantic content and appearing only in non-finite forms. I find this conceptually problematic.

Instead I will pursue the second option, which is a completely unified analysis of *gi-*, as this option is conceptually more attractive and more interesting to explore. There are two fundamental questions that a unified account raises. Firstly, there is the question of the special syntactic properties of *gi-*. If it is a verb, why doesn't it behave on a par with other complement-taking verbs? Secondly, there is the obvious question about the semantic content of *gi-*, namely, what it means to have lexical content in some uses but not in others. The answers that I give to these questions can be summarized as in (50). I will now discuss them in turn.

(50) a. *Gi-* is simultaneously a verb and a complementizer by way of lexicalizing V and C heads.
b. *Gi-* lacks any intrinsic lexical content and gets interpreted by a default interpretation rule.

4.2 Lexicalization of V and C

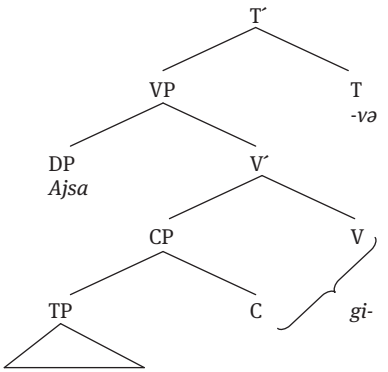
To account for the syntactic properties of the verb *gi-*, I would like to propose that *gi-* is a verb and a complementizer at the same time. I present a technical implementation of this proposal in a *nanosynactic* framework; see Starke (2009).

As in other syntactic approaches to morphology such as Distributed Morphology, in nanosyntax morphemes are syntactically represented. However, unlike Distributed Morphology nanosyntax assumes that morphemes can correspond to, or *spell out*, not only syntactic terminals but also larger syntactic structures. One option of such spellout is called *spanning*, a situation in which a morpheme realizes a continuous stretch of syntactic heads. To illustrate, the French portmanteau forms *du* and *au* can be analyzed as spelling out the preposition-determiner sequence, as in Svenonius (2012).

I would like to apply the same analysis to *gi-* and argue that *gi-* spells out the verbal head V and the head of its complement, that is, the complementizer head C.[23] For the uses of *gi-* as a contentful verb this is illustrated in (51b), where I present a partial tree corresponding to the sentence in (51a), repeated from above.

(51) a. *Ajsa* [*udlgo* *aavə-m* *ir-xə*] *gi-və.*
Ajsa soon grandfather-POSS.1SG come-PTCP.FUT say-PST
'Ajsa said my grandpa will come soon.'

b.

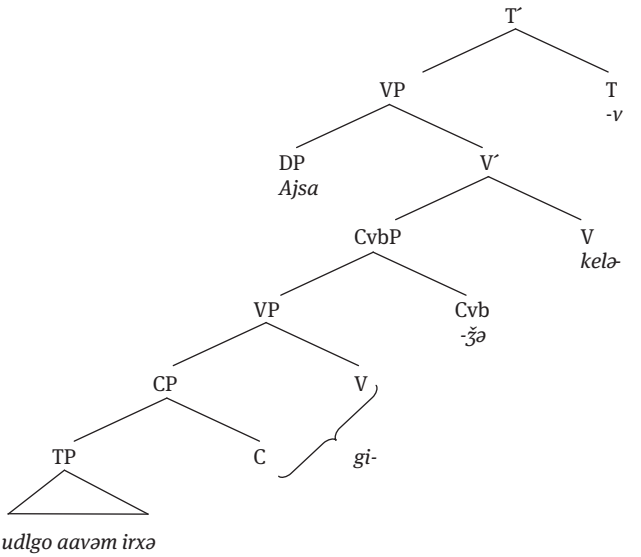


The analysis for the uses of *gi-* as a complementizer is illustrated in (52b), where I present a partial tree corresponding to the sentence in (52a), also repeated from above. Note that in line with syntactic approaches to morphology, I represent the converb suffix as a separate syntactic terminal (Cvb).

(52) a. *Ajsa* [*udlgo* *aavə-m* *ir-xə*] *gi-žə* *kelə-v.*
Ajsa soon grandfather-POSS.1SG come-PTCP.FUT say-CVB.IPFV tell-PST
'Ajsa said my grandpa will come soon.'

23 Svenonius (2012) observes that while spanning is usually confined to an extended projection of a single head, in certain cases a single morpheme appears to spell out a head (e.g., V) along with a head or heads in the extended projections of its complement. This would be exactly the case at hand.

b.



The analysis of *gi-* as lexicalizing both the V and C heads is able to account for why verbal morphology can attach to *gi-* and at the same time for why *gi-* in its uses as a lexical verb has special syntactic properties that set it apart from other matrix verbs. The fact that *gi-* is the only verb that is able to take sentential complements follows since *gi-* is also a complementizer and complementizers are closed-class elements. The adjacency requirement on the complement clause to *gi-* 'say' also follows from *gi-* being a complementizer since complementizers and clauses they introduce cannot be separated by material from the matrix clause. The incompatibility between the verb *gi-* and nominal complements also follows since C heads generally subcategorize for TP complements.

4.3 Default interpretation

In this section I will propose an account of the interpretation of *gi-*. I limit myself to a rather sketchy exposition leaving more detailed elaboration to future work. Under a unified analysis of *gi-* in its uses as a verb and its uses as a complementizer, we expect that it has a single interpretation. In principle there are two theoretical options. The first option is that *gi-* in both cases has lexical meaning and is interpreted as 'say'. This option is clearly not available since, as we saw in Section 3.2, *gi-* as a complementizer is not restricted to utterance predicates. Consequently, I will adopt the other option which is that *gi-* in both of its uses lacks lexical content. As to the apparent meaning of *gi-* in non-complementizer

sentences I would like to propose that this is the result of a rule of default semantic interpretation, which fills in the missing content, given in (53).

(53) A V with a propositional complement gets 'say' or 'believe' as its default semantic value (depending on the context).

The rule in (53) is essentially motivated by Pustejovsky (1995)'s analysis of cases like *John began a novel*. In Pustejovsky (1995)'s analysis the meaning of the understood verb (the action that John actually began) is recovered from the semantic content of the noun phrase complement, specifically, its qualia structure.[24] The understood verb in *John began a novel* tends to be interpreted as 'read' or 'write' since those actions are included in qualia structure of *novel* (as information related to how novels come about and what they exist for). The rule in (53) can be viewed as an application of this logic to propositions in general. The qualia structure of propositions would then include 'say' and 'believe' as the information about the origin and purpose of propositions, which generally come about by way of the speaker uttering them and are generally meant to be believed by the hearer.

Let's see how the rule in (53) operates in the environments where *gi-* appears. The crucial question will be how the choice is made between the 'say' and 'believe' interpretation. As to the complementizer uses of *gi-*, I will be assuming that the V associated with *gi-* forms a complex predicate with the higher verb so that the two verbal heads share argument structure. Now suppose that for the formation of a complex predicate, the thematic roles of the two verbs should match. Consequently, in the case of utterance predicates the agent (speaker) role should match that of the agent role of *gi-* and hence *gi-* should be interpreted as 'say' (as 'believe' lacks the agent role). In the case of propositional attitude predicates, commentative predicates and other individual predicates such as *nääl-* 'hope', *soŋs-* 'hear', *med-* 'know' the role of the holder (of the propositional attitude) should match that of the holder role of *gi-* and hence *gi-* should be interpreted as 'believe'. Note that this exactly predicts the non-factive interpretation of the sentences involving *med-* 'know' and commentative predicates such as *bajrəl-* 'be glad', as I discussed in Section 3.2. Moreover, this might also explain why the complementizer *giǯə* is incompatible with factive verbs like *mart-* 'forget' and *üz-* 'see' if these verbs cannot be coerced into a non-factive interpretation.[25]

24 The qualia structure is part of the lexical meaning of a word embracing information about the distinguishing features (so-called *formal* role), parts (*constitutive* role), purpose (*telic* role) and origin (*agentive* role) of an object designated by this word.

25 In the case of the complementizer *gisən* restricted to complements of nouns (see Section 3.1; the interpretation of *gi-* should also be controlled by the content of the nominal predicate (pre-

As to what I called the contentful uses of *gi-*, I would like to propose that here pragmatic considerations favor the 'say' interpretation over that of 'believe'.[26] One reason for this might be the following. Suppose that the meaning of the verb 'say' includes 'believe' as its component. This would be the case if speech acts are viewed as external manifestations of mental states.[27] Consequently, 'say' would be more informative than 'believe'. As a result, we would expect that 'say' would be preferred over 'believe' as the more informative interpretation essentially in a Gricean fashion. Clearly, pragmatic consideration such as these will only come into play in the absence of complex predicate formation. This will ensure that in the complementizer uses *gi-* does not always come out as 'believe' but rather is interpreted depending on the content of the higher verb whereas in the unembedded uses the interpretation of *gi-* is fixed to 'say'.

5 Conclusion

In this paper I have looked at the verb *gi-* 'say' and its non-finite forms (mostly *giǯə* but also *giʁäd* and *gisən*), which function as complementizers. I have shown that the verb *gi-* 'say' has peculiar syntactic properties shared with complementizers derived from *gi-* 'say'. I have examined the distribution of the complementizers showing that they are compatible with a wide range of complement-taking verbs going beyond utterance predicates provided the context is non-factive (and also non-assertive in the case of *giʁäd*). I have also shown that the complementizers distributionally pattern with the respective non-finite forms (complements of nouns requiring the form *gisən*).

I have proposed a unified analysis for the verb *gi-* and its forms used as complementizers according to which *gi-* is simultaneously a verb and a complementizer (by way of embracing both V and C heads in a nanosyntactic fashion) in both of its uses as a verb and a complementizer. The analysis is able to account for the peculiar syntactic properties of the verb *gi-* and the distributional pattern displayed by the complementizers, which would be totally unexplained if the verb *gi-* and the complementizers derived from it were (synchronically) independent lexical items.

---

sumably in a different way than complex predicate formation). However, for reasons of space I have to leave the discussion of these cases for future research.

26 I am grateful to a reviewer for the bringing up this issue.

27 See Shinzato (2004) for a defense of this view.

I have also argued that *gi-* lacks any intrinsic lexical content even when used as a verb and gets interpreted (in those cases) by a default interpretation mechanism, which fills in the missing content. This completely unified analysis of *gi-* can be viewed as a conceptually simpler alternative to a more obvious grammaticalization analysis, according to which the verb *gi-* present in the complementizers has lost its lexical meaning despite retaining some of the verbal properties.

References

Baranova, Vlada V. 2010. Puti grammatikalizacii glagola gi- 'govorit' v kalmyckom jazyke [Paths of grammaticalization of the verb gi- 'say' in Kalmyk]. In Zoja M. Šaljapina (ed.), *Trudy mežinstitutskoj naučnoj konferencii "Vostokovednye čtenija 2008"* (Moskva, 8–10 oktjabrja 2008 g.) [Proceedings of the multi-institutional conference "Readings in Oriental Studies 2008", Moscow, October 8–10, 2008], 96–114. Moscow: IV RAN.

Blaesing, Uwe. 2003. Kalmuck. In Juha Janhunen (ed.), *The Mongolic languages*, 229–248. London: Routledge.

Haegeman, Liliane & Barbara Ürögdi. 2010. Referential CPs and DPs: An operator movement account. *Theoretical Linguistics* 36. 111–152.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Higginbotham, James. 1983. The logic of perceptual reports: An extensional alternative to situation semantics. *The Journal of Philosophy* 80. 100–127.

Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4. 465–497.

Kiparsky, Carol & Paul Kiparsky. 1970. Fact. In Manfred Bierwisch & Karl Erich Heidolph (eds.), *Progress in Linguistics*, 143–173. The Hague: Mouton.

Knyazev, Mikhail Yu. 2009. Sentencial'nye dopolnenija v kalmyckom jazyke [Sentential complements in Kalmyk]. In Sergey S. Say, Vlada V. Baranova, and Natalia V. Serdobolskaya (eds.), *Issledovanija po grammatike kalmyckogo jazyka* [Studies in Kalmyk grammar], 525–580. Saint Petersburg: Nauka.

Lehmann, Christian. 1995. *Thoughts on grammaticalization*. München: Lincom Europa.

Mittwoch, Anita. 1990. On the distribution of bare infinitive complements in English. *Journal of Linguistics* 26. 103–131.

Noonan, Michael. 1985. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol 2, *Complex constructions*, 42–140. Cambridge: Cambridge University Press.

Prokhorov, Kirill. 2009. Kalmyckie formy kosvennyx naklonenij: morfologija, semantika, sintaksis [Oblique moods in Kalmyk: Morphology, semantics, syntax]. In Sergey S. Say, Vlada V. Baranova, and Natalia V. Serdobolskaya (eds.), *Issledovanija po grammatike kalmyckogo jazyka* [Studies in Kalmyk grammar], 160–224. Saint Petersburg: Nauka.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.

Say, Sergey S. 2009. Argumentnaja struktura kalmyckix kauzativnyx konstrukcij [Argument structure of causative constructions in Kalmyk]. In Sergey S. Say, Vlada V. Baranova, and

Natalia V. Serdobolskaya (eds.), *Issledovanija po grammatike kalmyckogo jazyka* [Studies in Kalmyk grammar], 387–464. Saint Petersburg: Nauka.
Say, Sergey S., Vlada V. Baranova & Natalia V. Serdobolskaya (eds.). 2009. *Issledovanija po grammatike kalmyckogo jazyka* [Studies in Kalmyk grammar]. Saint Petersburg: Nauka.
Sheehan, Michelle & Wolfram Hinzen. 2011. Moving towards the edge. *Linguistic Analysis* 37. 405–458.
Shinzato, Rumiko. 2004. Some observations concerning mental verbs and speech act verbs. *Journal of Pragmatics* 36. 861–882.
Shklovsky, Kirill & Yasutada Sudo. 2013. Indexical Shifting in Uyghur and the Syntax of Monsters. In Seda Kan, Claire Moore-Cantwell & Robert Staubs (eds.). *NELS 40: Proceedings of the Fortieth Annual Meeting of the North East Linguistic Society*, 151–166. Amherst: Graduate Linguistics Student Association.
Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. [Special issue]. *Nordlyd* 36(1).1–6. septentrio.uit.no/index.php/nordlyd/article/download/213/205 (accessed 3 March 2014).
Svenonius, Peter. 2012. *Spanning*. Manuscript, University of Tromsø. ling.auf.net/lingbuzz/001501/current.pdf (accessed 3 March 2014).

Natalia Serdobolskaya

Semantics of complementation in Adyghe

1 Introduction

This paper investigates the semantic distribution of complement constructions in the Adyghe language (North-West Caucasian). The complementation system of Adyghe presents many interesting issues that are not widely attested typologically. An interesting point is the possibility of unsubcategorized case marking in complementation: clausal arguments show variation in case marking with one and the same complement-taking verb, a phenomenon which is not attested with nominal arguments. Another issue is the use of the same complement construction to encode events with emotive verbs (like *I like dancing*) and false propositions (like *he had been to China* in *He said he had been to China, but I know he was lying*). Meanwhile there is a special "factive form" used to encode facts and manner complements.

Complementation in Adyghe has previously been described in both reference grammars and the specialized literature on subordination (see Zekokh 1976; Kerasheva 1984 and others). However, these works do not focus on the semantic types of complement clauses. I consider the main complementation strategies in Adyghe in terms of semantic notions including fact vs. event vs. proposition, presupposition vs. assertion, and epistemic meaning. In particular, competition between the various complementation strategies is dealt with in detail. An overview of the relevant parameters, and the tests that help to differentiate between the semantic types of complement clauses, are given in Serdobolskaya (this volume), which is focused on Ossetic. A list of the complement-taking predicates considered here is given in the Appendix.

Most of the Adyghe data presented here were collected by elicitation and from texts recorded during fieldwork in 2004, 2005, 2006 and 2008 in the village of Khakurinokhabl, in the course of field trips carried out by the Russian State University for the Humanities (led by Jakov Testelec, Nina Sumbatova, and Svetlana Toldova). The elicited examples are given without reference, while examples taken from recorded oral texts are labelled TEXT.

I will use the following terms: *complement-taking predicate* – a predicate that can take clausal (and potentially also nominal) arguments; *matrix clause* – the clause headed by the complement-taking predicate; *complement clause* – a clausal argument (marked with square brackets); *dependent/embedded clause* – any type of subordinate clause; *complementation strategy* – the morphosyntactic construction used for the complement clause.

The paper is structured as follows. Section 2 presents a list of complementation strategies in Adyghe. Section 3 is dedicated to the morphosyntactic properties of the main complementation strategies. In Section 4, I analyze the semantics of the different complementation strategies and explain the choice of construction in those cases where competition exists.

2 Adyghe: typological features and main complementation strategies

2.1 Typological features of Adyghe

Adyghe (West Circassian) belongs to the Abkhazo-Adygh branch of the North-Caucasian language family; it is spoken in the Republic of Adygea and in the Krasnodarskij district in Russia, and also in Turkey, Syria, Israel and several other countries. The total number of speakers does not exceed 500,000. Standard Adyghe is based on the Kemirgoy dialect spoken in Adygea. The idiom of the village of Khakurinokhabl, described in this paper, is largely Kemirgoy, with a very small number of features (mostly phonetic) belonging instead to the Abadzeh dialect.

The Adyghe language is characterized by a large consonant system, whereas only three vowels are distinguished. The main typological features of Adyghe are ergative alignment, polysynthesis in the verb morphology, a small number of nominal cases, free word order, and somewhat blurred morphological distinctions between the parts of speech (cf. Rogava & Kerasheva 1966; a grammatical sketch is provided by Arkadiev et al. 2009).

Adyghe has four morphological cases: oblique (ergative), absolutive (which marks patient-like arguments of transitive verbs (1) and the single argument of intransitive verbs (2), instrumental and adverbial. The oblique case is used to mark agent-like arguments of transitive verbs (1), indirect objects with various semantic roles, and adjuncts (cf. *šə-m* and *zawe-m* in (2)). This case is termed oblique here on the basis of its functional range (for further arguments against the term ergative, cf. Letuchiy 2012).

(1) *haç̌’e-**m*** *ç̌’ale-**r*** *ə-ɬeʁʷə-ʁ*.
guest-**OBL** boy-**ABS** 3SG.A-see-PST
‘The guest has seen the boy.’
(Rogava & Kerasheva 1966: 65)

(2) *ç'ale-**r*** *šə-**m*** *tje-s-ew* *zawe-**m*** *ḳ^{w}a-ʁe.*
boy-**ABS** horse-**OBL** LOC-sit-ADV war-**OBL** go-PST
'The boy went to war on horseback.'
(Rogava & Kerasheva 1966: 65)

The instrumental case encodes a large number of meanings, such as instrument, means, measure, and motion through an area (prolative) (Serdobolskaya 2011). The adverbial case is mostly used to mark adverbs, secondary predication, and converbs, as well as in some other functions (see Section 3.5 for details). The absolutive and the oblique case mark core arguments of the verb, as opposed to the remaining two cases, which are most often used to introduce adjuncts. Henceforth I will thus use the term *core cases* for the oblique and the absolutive, and *non-core cases* for the instrumental and the adverbial.

In Adyghe, most semantic roles are encoded by verbal prefixes. For example, the beneficiary is introduced by a special prefix on the verb:

(3) *se* *txələ-xe-r* *sə-š'efə-ʁe-x.*
I book-PL-ABS 1SG.A-buy-PST-PL
'I bought the books'.
(Letuchiy 2009: 331)

(4) *se* *ç'ale-xe-**m*** *txələ-xe-r* *a-**fe**-s-š'efə-ʁe-x.*
I boy-PL-**OBL** book-PL-ABS 3PL.IO-**BEN**-1SG.A-buy-PST-PL
'I bought the books for the boys'.
(Letuchiy 2009: 331)

In (4) the noun phrase denoting the beneficiary is marked with the oblique case, the same case marker that is used to encode agent-like arguments of transitive verbs. The valency-increasing prefix *fe-* on the verb specifies its role in the sentence as a beneficiary (in reference grammars of Adyghe this prefix is termed "version" [Rogava & Kerasheva 1966]).

The Adyghe sentence can host several noun phrases marked with the oblique case, whose semantic roles are specified by prefixes on the verb. These prefixes are marked for the person and number of the corresponding participants. Meanwhile, a clause may contain only one absolutive noun phrase. The overwhelming majority of verbs require an absolutive argument.

The tense system of Adyghe includes present tense (no special morphological marker; however, a "dynamicity" morpheme attaches to verbs, except for stative predicates), future tense (the marker *-š't*), past tense (the marker *-ʁ*), the remote past in *-ʁaʁ* (PST-PST) and the imperfect in *-š'tə-ʁe* (the past tense of the auxiliary, which is homonymous with the future tense; for the arguments against the uniform analysis cf. footnote 5).

2.2 Complementation strategies in Adyghe

Adyghe represents an interesting case for the typology of complementation, since in this language the case markers develop a complementizing function when used with verbal stems. The possibility of choosing more than one case marker in a given context multiplies the number of complementation strategies available, as will be shown in the following sections.

The main complementation strategies in Adyghe (first described in Gerassimov 2006) are the following: verbal stem without tense suffixes or subordination markers (5), potentialis without case markers (6), factive form in *zere-*[1] (7), verbal stem with case markers (8) and potentialis with case markers (9).

(5) *ʔʷənč'əbze-xe-r* ***z-ʁe-ḳʷed*** *s-jə-xabz.*
key-PL-ABS **1SG.A-CAUS-lose** 1SG.PR-POSS-law
'I sometimes lose the keys (lit. It's my law to lose the keys).'

(6) *laʁe-xe-r* *s-thač'ə-**n*** *faje.*
plate-PL-ABS 1SG.A-wash-**POT** must
'I must wash the dishes.'

– factive form with the prefix *zere-* (the factive form takes the case marker required by the complement-taking predicate):

(7) [*aχš'e* *deʁʷ* *qə-**zer**-jə-ʁaχe-re-**r***] *s-əgʷ* *r-j-e-hə.*
money good DIR-**FCT**-3SG.A-earn-DYN-**ABS** 1SG.PR-heart LOC-3SG.A-DYN-carry
'I'm pleased that he earns a lot.'

– verbal stem with tense markers (the zero marker of the present tense in [8]) and with the case markers of adverbial or instrumental case:

(8) [*azemat* *qa-ḳʷ-**ew***] *qə-s-š'-eχʷə.*
Azamat DIR-go-**ADV** DIR-1SG.IO-LOC-DYN-become
'It seems to me that Azamat will come.'

1 Reference grammars use different terms for these forms. The verbal stem without suffixes is labelled "bare stem" in Kerasheva (1984: 121–134), "non-finite form without suffixes" in Rogava & Kerasheva (1966: 329), and "infinitive" in Kumakhov (1989: 277–278). The potentialis is often called "masdar" (Rogava & Kerasheva 1966; Kumakhov 1989; Zekokh 2002), but the term "action nominal" is used in Paris & Batouka (2005). Kumakhov (1989) uses the term "infinitive" for the potentialis without case markers, and "masdar" for the potentialis with case markers. The masdar (i.e. potentialis) with the adverbial case marker in particular is described as the "supine" in Rogava & Kerasheva (1966) and Kerasheva (1984), while other researchers retain the term "masdar" here. The factive form is referred to as the "participle" (Rogava & Kerasheva 1966: 111), and the verbal stem with the adverbial case as the "converb"; there is no special term for the

– potentialis with case markers:

(9) [*sə-tje-fe-**n**-**ç'e***] *s-e-š'əne.*
1SG.ABS-LOC-fall-**POT-INS** 1SG.ABS-DYN-fear
'I'm afraid of falling down.'

I consider this a distinct construction from the potentialis without case markers, because the two differ in their morphosyntactic properties and syntactic distribution. The potentialis without case markers is used with modal, phasal, and aspectual predicates, i.e. predicates that do not denote an independent situation, but modify another situation (in terms of phase, modality, or aspect respectively). Typologically it is known that such verbs often head clause union structures with their dependent verbs (cf. Aissen 1974; Noonan 1985: 138–142). According to Kimmelmann (2010), in Standard Adyghe most of these verbs form monoclausal constructions with their "complements".

In contrast, the potentialis with case markers is used with verbs that denote an independent situation, such as mental/emotive verbs and verbs of speech and perception. This construction will be of major concern in this paper, as will the factive form (7) and the verbal stem (8). The first two strategies, exemplified in (5)[2] and (6), will not be considered.

– indirect question strategies

Complement clauses with general questions take the conditional marker *-me* with the additive particle *-jə* (10). The particle is optional in such contexts. In conditional clauses in Adyghe, the marker *-me* introduces a real/unreal/counterfactual condition, and the combination of the markers *-me* and *-jə* denotes concession.

(10) *ç'ale-m* *ə-ŝe-r-ep* [*pŝaŝe-r* *me-čəje-**m-jə***].
boy-OBL 3SG.A-know-DYN-NEG girl-ABS DYN-sleep-**COND-ADD**
'The boy does not know if the girl is sleeping.'
(Arkadiev et al. 2009: 96)

Indirect wh-questions are introduced by relative clause constructions, where the question focus is the head, e.g. (11) literally 'I do not know the one where Aslan has gone'.

verbal stem with the instrumental case. The rationale for adopting the terms used in this paper are discussed in detail in Serdobolskaya (2009a) and (2009b).

2 The construction in (8) should not be confused with that seen in (5): in (5), the "bare" verbal stem is used, i.e. without any tense suffixes or subordination markers, while in (8) the verbal stem takes both tense suffixes and case markers, which serve as subordinators.

(11) *se* [*asλan* ***zə-de***-*ḳʷa-ʁe-**r***] *s-ŝe-r-ep*.
I Aslan **REL.IO**-LOC-go-PST-**ABS** 1SG.A-know-DYN-NEG
'I do not know where Aslan has gone.'
(Arkadiev et al. 2009: 93)

I do not consider these constructions in detail, since they have a very special distribution, i.e. they only appear in the context of indirect questions.

A number of other complementation strategies exist which are used more rarely than those enumerated in (5)–(9). Since their use is limited to a small number of constructions/complement-taking predicates, I describe them very briefly here.

– verbal noun in *-č'e*:

The verbal noun has the meaning of manner, e.g. *ḳʷač'e* 'gait, step' from *ḳʷe-* 'go', *š'əʔač'e* 'way of life' from *š'əʔe-* 'live' (Kumakhov 1964: 121), and it heads sentential complements with manner meaning:

(12) [*a-xe-m* *ja-qe-ŝʷa-**č'e***] *s-λeʁʷə-ʁe*.
DEM-PL-OBL 3PL.PR+POSS-DIR-dance-**NMLZ** 1SG.A-see-PST
'I saw how they were dancing.'

Unlike the verbal stem with the instrumental, the verbal noun in *-č'e* shows nominal morphosyntactic properties, e.g., it takes the possessive marker and cannot host all the arguments of the initial verb. It also differs from this form morphonologically: unlike the instrumental case marker, it triggers the *e* → *a* alternation in the preceding morpheme.

– the grammaticalized form of the verb of speech *ʔʷe-* 'say' with the additive particle *-jə*:

(13) [*"se* *prosto* *sə-b-de-semerqewə-ʁ* *nah-č'e"* ***ə-ʔʷ-jə***]
I just 1SG.ABS-2SG.IO-COM-joke-PST than-INS **3SG.A-say-ADD**
qə-s-jə-ʔʷa-ʁ.
DIR-1SG.IO-3SG.A-say-PST
'"I have only played a practical joke on you, nothing else", he said to me.'
(Text)

This form is grammaticalized and has somewhat "bleached" semantics, since it can also mark citations with predicates that do not necessarily presuppose speech (14) and can be used to mark complements of the same complement-taking predicate *ʔʷe-* 'say' (13) without giving rise to tautology. However, it preserves agreement with the agent of the matrix clause, cf. the 1st person marker in (14), contrasting with the 3rd person in (13).

(14) *jeṭane direktorə-m thape-m fə-tje-stx-jə* [*“ḳʷe*
then school.head-OBL paper-OBL BEN-LOC-1SG.A-write-ADD go.IMP
čʼe-šʼ je obed gʷere j-e-ʁe-ʂe je nemәčʼ gʷere
LOC-take.away.IMP or lunch some OBL-DYN-CAUS-do.IMP or other some
fe-ʂ̣” ***s-ʔʷ-jə***].
BEN-do.IMP **1SG.A-say-ADD**
‘...then I wrote a note to the school head (lit. saying): “Come and take her [the inspector] to lunch or whatever, or do something else with her”.’
(Text)

– parataxis

This strategy is distinct from that using the verbal stem (8) in terms of its morphology. Firstly, it does not take case markers; secondly, there is a clear indicator of the non-subordinate status of the semantically dependent clause, namely the “non-subordinate” negation *-ep* (see Section 3.1 for the negation in Adyghe):

(15) *se* [*azemat qe-ḳʷe-šʼt-**ep***] *s-ʂʷeʂə.*
I Azamat DIR-go-FUT-**NEG** 1SG.A-think
‘I think that Azamat is not coming.’

This type of construction has only been observed with non-factive mental verbs of opinion. However, the distribution of this strategy has not been investigated in detail.

– temporal and conditional markers:

(16) a. *saše me-šʼəne* [*ʂʷənčə **zə-χʷə-čʼe***].
Sasha DYN-fear dark TEMP-become-INS
b. *saše me-šʼəne* [*ʂʷənčə χʷə-**me***].
Sasha DYN-fear dark become-COND
‘Sasha is afraid of the dark.’

These markers are used in complements denoting generic events, as in (16), with emotive and evaluation predicates and the verb *ježe-* ‘wait’.

The problem with these constructions is whether they should be classified as argument or adverbial clauses. Semantically the conditional/temporal clauses in (16) are filling the valency slot of the matrix predicate, since the stimulus of emotion belongs to the set of arguments of an emotive verb. However, they are encoded with means that are otherwise found in head conditional/temporal adverbial clauses.

One test for the argument status of a noun phrase (suggested in Apresyan 1974) concerns the possibility of omitting the noun phrase in question[3]. On the

3 It has been pointed out in the literature that this criterion does not always give plausible results and must be used with caution in relation to labile verbs (cf. the possibility of *She loves* in English

basis of this criterion, the clauses in (16) are to be classified as adjuncts. They can be omitted without leading to ungrammaticality, e.g. it is possible to say *s-e-š'əne!* [1SG.ABS-DYN-fear] 'I am afraid' without expressing the stimulus. By contrast, the stimulus cannot be omitted with the verb *š'eš'əne-* 'fear', which contains the preverb *š'e-* (note that with this verb the function of the preverb *š'e-* is to introduce a new valency slot). The same is true for other complement-taking predicates that take conditional/temporal clauses. Thus, these constructions do not meet the syntactic criterion defining complementation, and are not considered in this paper.

It is significant that in Adyghe coreferentality between the arguments of the matrix and the dependent clause does not play a role in the choice of complementation strategy. Although there is a form traditionally termed the "infinitive" (potentialis without a case marker; in other reference works, this is taken as the verbal stem without suffixal markers), it is reserved for phasal, aspectual and modal complement-taking predicates only. Other classes of verbs do not show a preference for any of the analyzed strategies on the basis of the corefentiality pattern (cf. [20] and [50]) with the complement-taking predicate 'fear', or (52) and (83a–b). with 'hope').

In the next section I address the morphosyntactic features of the main complementation strategies in Adyghe.

3 Morphosyntactic properties of complement clauses

In this section, I consider the morphosyntax of complement clauses headed by the potentialis, the verbal stem, and the factive form. All three forms take case markers (and postpositions) that mark their subordinate status.

Adyghe predominantly uses non-finite clauses in complementation, as in other types of subordinate constructions. The question arises as to the "degree" of finiteness of subordinate clauses. A number of parameters have been suggested in works on (non-)finiteness, such as tense/aspect/mood marking and the presence of nominal cases (Lehmann 1988; Givón 1990; Malchukov 2004; Kalinina 2002). The following morphosyntactic parameters are relevant for Adyghe:

---

without the noun phrase referring to the stimulus). In Adyghe, arguments of two-place verbs can also be omitted under certain semantic conditions; however, this does not apply in the case of many verbs with valency-increasing prefixes, including *š'e-š'əne-*.

– Morphological properties of verbal forms
  Verbal: mood and modality, tense, aspect, valency-changing markers, negation, agreement markers, dynamicity marker[4];
  Nominal: presence of an overt nominalizer, possibility of taking case markers (and/or articles), nominal number and possessive markers.
– Syntactic properties of verbal forms
  Marking of the core arguments (if this reflects the pattern used in independent clauses or in noun phrases), (in)compatibility with postpositions, (in) compatibility with adverbs/adjectives, (in)ability to head an independent clause.

3.1 Morphosyntax of complement clauses headed by the potentialis

The potentialis can take nearly all the verbal grammatical markers; see example (17) with the 2nd person marker *-p-*, (19) with the number agreement suffix, (17) with the valency-increasing prefix *fe-* and (17)–(18) with the aspectual markers (the preverb *qə-* [17] and the refactive suffix *-ž'ə* [18]).

(17) [*tawrəχə-r* *qə-**p-fe-s**-ʔʷete-**nə**-m* *əpe*]
tale-ABS DIR-**2SG.IO-BEN-1SG.A**-tell-**POT**-OBL before
wə-z-ʁe-šxe-š'ət.
2SG.ABS-1SG.A-CAUS-feed(ANTIP)-FUT
'I'm going to feed you before I tell you a fairy tale.'

(18) *maše* *ja-dež'* *ḳʷe-**ž'ə-n**-ew* *z-je-ʁe-hazərə.*
Masha 3SG.IO-towards go-**RE-POT**-ADV REFL.ABS-DYN-CAUS-ready
'Masha plans to return home (lit. to her).'

There are two negation markers in Adyghe: the prefix *mə-* is used in subordinate clauses, while the suffix *-ep* marks the head of an independent clause (e.g. *ḳʷe-r-**ep*** [go-DYN-NEG] 'he doesn't go' vs. *ə-**mə**-ləʁʷə-ʁe jə-ŝawe-m* [3SG.A-NEG-see-PST POSS-boy-OBL] 'his son that he has never seen' (Arkadiev 2009 et al.: 45). The potentialis in complement clauses takes only the non-finite negation, as in all subordinate clauses:

(19) [*we* *laʁe-xe-r* *wə-**mə**-thaç'ə-n-x-ew*] *sə-faj.*
you plate-PL-ABS 2SG.A-**NEG**-wash-POT-PL-ADV 1SG.ABS-want
'I don't want you to wash the plates (lit. I want that you not wash the plates).'

4 In Adyghe, dynamic verbs in the present indicative take one of the dynamicity markers *ma-/me-*, *-e-*, or *-re-* (for the exact distribution see Rogava & Kerasheva 1966).

Note than in independent clauses the potentialis takes the finite negation, as shown in (23).

Past tense markers are not in complementary distribution with the potentialis; they can co-occur in one and the same verbal form:

(20) [*vaze* *qʷətə-**ʁe-n**-ç'e*] *s-e-š'ane*.
vase break-**PST-POT**-INS 1SG.ABS-DYN-fear
'I fear that the vase has broken (during transportation).'

The future tense marker *-š't* cannot occur in the potentialis form. However, it is unclear whether this prohibition is due to the non-finite properties of the potentialis or whether it can be explained by the semantics of the two forms. The future tense and the potentialis are semantically similar, both signalling the future time reference of the situation (whereas the potentialis has peculiar semantic nuances, which are discussed below in this section). Hence, it can be assumed that the two markers are incompatible in a single verbal form because of the tautology effect. The two markers cannot co-occur in verbal forms that head independent sentences either.[5]

The mood and dynamicity markers are incompatible with the potentialis, not only in complement clauses, but also in the head of independent sentences. The markers of the conditional in *-me*, concessional in *-mjə* and desiderative in *-ʁʷet* can co-occur with the future tense marker *-š't*, but they are incompatible with the potentialis (Rogava & Kerasheva 1966: 192–201). This may serve as an argument for the non-finiteness of the complement clauses with the potentialis, as well as in independent clauses.

Hence, from the point of view of verbal morphological categories, the potentialis demonstrates nearly all the properties that are characteristic for verb forms that head independent sentences, except for the ability to take markers of non-indicative mood and dynamicity.

The potentialis also demonstrates a number of nominal features. The marker itself can be analyzed as the nominalizer; note that most reference grammars label this form the "masdar" (Zekokh 2002; Kerasheva 1984; Kumakhov 1989;

5 There are verbal forms that seem to violate this restriction, e.g. *ma-ḳʷe-**š'tə**-n* [DYN-go-**AUX-POT**] 'he probably goes' (Arkadiev et al. 2009: 48). However, these forms are interpreted by Arkadiev et al. (2009) as the potentialis forms of the auxiliary verb *š'tə-*, which encodes epistemic modality. This auxiliary also has other tense forms, such as *ə-ʔʷe-**š'tə-ʁa-ʁe*** [3SG.A-say-**AUX-PST-PST**] 'he would say'. The argument for differentiating this auxiliary from the future tense marker is the position of the negation marker *-ep*: with the future marker it is postposed to all other suffixes (including the future tense marker *-š't*), while in the construction with the auxiliary it precedes the auxiliary marker *-s't*: for example, *š'ə-sə-ʁ-ep-š'tə-n* [LOC-sit-PST-NEG-AUX-POT] 'he probably didn't sit'.

Rogava & Kerasheva 1966) or "action nominal" (Paris and Batouka 2005). It is not easy to say whether this marker should be viewed as a subordinate clause marker or a modal marker. The potentialis is most often found in subordinate clauses that have future temporal reference or gnomic meaning. However, it can also head independent sentences:

(21) *se tort s-ʂə-š'ɬ/-**n**.*
I cake 1SG.A-do-**FUT/-POT**
'I'll bake a cake.'

(22) *sə-qə-b-de-ʔepəʔa-**n**-a?*
1SG.ABS-DIR-2SG.IO-COM-help-**POT**-Q
'Should I help you?'

(23) *laʁe-r s-ʔepəzə-**n**-ep.*
dish-ABS 1SG.A-drop-**POT**-NEG
'{I'll be very careful and} I won't drop the dishes, {don't worry}.'

The acceptability of this form in independent clauses cannot, however, be taken as a good argument against analyzing it as a subordinate clause marker. The potentialis in independent clauses cannot take indirect mood and dynamicity markers. It has been demonstrated in Kalinina (2002), Evans (2007) and other works that subordinate verbal forms can occur as the head of independent sentences as a result of so-called *insubordination*. Analysis of the contexts where the potentialis can head an independent sentence shows the following (cf. Serdobolskaya 2009a for details): the potentialis is used if the independent clause has reference to the future and does not constitute the main assertion of the sentence. Let us clarify this point. In (21) the potentialis introduces the semantic nuance of spontaneous decision; in (22) it encodes deontic necessity; and (23) it suggests an attempt at persuasion. Native speakers comment that the potentialis is possible if the speaker is *assuring* the hearer that he will not drop the dishes, rather than simply stating it as a fact. Such "fuzzy" explanations have one point in common: native speakers always try to find some semantic matrix predicate (and a matrix clause) that is not expressed in the sentence headed by the potentialis: in (21) it is 'I have just decided', in (22) it is a marker of deontic modality, in (23) it is 'I assure you'. Such a semantic predicate, which is absent in the syntactic structure, is present in the insubordinated use of infinitives, participles and other verbal forms that head non-finite clauses (Evans 2007). Hence, the marker of the potentialis can be analyzed as a marker of subordination with a semantic matrix predicate, and the examples with the potentialis in independent clauses can be interpreted as a result of insubordination. Thus, the potentialis can be described semantically as a form mostly referring to the future, and having an "insubordinated" use in

main clauses, i.e. when found in main clauses it does not itself make a semantic assertion but requires a semantic matrix predicate, cf. Evans (2007) for the types of semantic matrix predicates found in insubordination.

The nominal morphological features of the potentialis are the following. First, the potentialis can take all four nominal cases, cf. (24)–(26) and (30). (The nominal number affix *-xe* also functions as a verbal plural agreement marker (24), and thus cannot be used as an argument regarding the finiteness of a construction.)

(24) [*zeč'e* *çəf-xe-r* *nasəpə* *ŝə-**n-x**-ew*] *tə-faj.*
all person-PL-ABS happy do-**POT-PL**-ABS 1PL.ABS-want
'We want everybody to be happy.'

Second, the person and number agreement prefixes can be omitted on the potentialis (25), while this is unacceptable in verbal forms that head independent sentences.

(25) [*mašine* *(**s**)-fə-nə-r*] *qə-s-fe-qjən.*
car (**1SG.A**)-drive-POT-ABS DIR-1SG.IO-BEN-hard
'It is hard for me to drive a car.'

The omission of agreement markers is only possible with the potentialis in core cases; it is not possible in non-core cases, cf.:

(26) a. [*mašine-r* ***s**-fə-n-č'e/ew*] *qə-s-fe-qjən.*
car-ABS **1SG.A**-drive-POT-INS/ADV DIR-1SG.IO-BEN-hard
b. *[*mašine-r* *fə-n-č'e/ew*] *qə-s-fe-qjən.*
car-ABS drive-POT-INS/-ADV DIR-1SG.IO-BEN-hard
'It is hard for me to drive the car.'

It can be concluded that the potentialis forms with and without agreement markers head different constructions in Adyghe. The construction without the markers shows a higher degree of nominalization than the construction with the agreement markers.

Third, the potentialis can take the nominal marker of possession *jə-*. This prefix occurs postposed to the person and number prefix and marks alienable possession on nouns, as in (27). It can occur on the potentialis (28).

(27) a. *s-šə*
1SG.PR-brother
'my brother'
b. *s-**jə**-š*
1SG.PR-**POSS**-horse
'my horse'
(Arkadiev et al. 2009: 56)

(28) *s-**jə**-ḳ^{w}e-n* *deʁ^{w}ə* *s-**jə**-če-n* *nah.*
1SG.PR-POSS-go-POT good 1SG.PR-**POSS**-run-POT than
'I walk well, but I can hardly run (lit. My walking is good, compared to my running).'

Such examples are possible if the potentialis denotes an activity with gnomic meaning or a cultural event, i.e. in more nominalized contexts. Complement clauses that have concrete temporal reference do not allow such a construction. Even with a gnomic interpretation the possessive markers are not always possible:

(29) a. **[ç'ale-m* ***jə**-qe-ŝ^{w}e-n*][6] *jə-ç'as.*
boy-OBL **POSS**-DIR-dance-POT POSS-favourite
b. [*ç'ale-m* *qe-ŝ^{w}e-nə-r*] *jə-ç'as.*
boy-OBL DIR-dance-POT-ABS POSS-favourite
'The boy likes dancing.'

I suggest that (28) shows a higher degree of nominalization than the form in (29b).

Different syntactic structures must therefore be postulated for (28) and (29b). Hence, the potentialis in Adyghe can head a number of constructions that exhibit different degrees of nominalization (and desententialization, cf. Lehmann 1988);[7] see Ershova (forthcoming) for a similar analysis of Kabardian constructions. The least nominalized constructions are those that include the potentialis with adverbial/instrumental markers. The highest degree of nominalization is observed with forms demonstrating a clear semantic shift towards the denotation of activities or cultural events (or even objects, such as *šxən* 'food' from the verb *šxə-* 'eat') rather than propositions or events. This semantic distinction is common for many languages that possess various syntactic types of nominalization.

In terms of its syntactic behaviour, the potentialis mostly demonstrates verbal properties: its core arguments are marked according to the pattern used in independent finite clauses (e.g. absolutive in [25]), and adjuncts are most often marked as adverbs (30). In Adyghe, adverbs derived from adjectives are marked with the adverbial marker *-ew*, e.g. *ž'ə* 'early' (adjective), *ž'ew* 'early' (adverb).

6 I do not give an asterisked example with the possessive on the absolutive form of the potentialis, since the possessive marker is incompatible with nouns in the absolutive (Rogava & Kerasheva 1966: 62–64).

7 It is well known that in many languages one and the same nominalized form can head constructions with different degrees of nominalization; see the analysis of the constructions headed by the English gerund *Mary's singing of the song* (nominal pattern), *Mary singing the song* (verbal pattern), *Mary's singing the song* (mixed pattern) in Abney (1987).

(30) [***ž'-ew*** *sə-teǯ'ə-n-əm*] *s-jezeš'ə-ʁ.*
early-ADV 1SG.ABS-get.up-POT-OBL 1SG.ABS-fed.up-PST
'I'm fed up with getting up early.'

(31) *sə-g^{w}* *r-jə-hə-ʁ* ***təʁ^{w}ese-re*** *qe.ŝ^{w}e-nə-r.*
1SG.PR-heart LOC-3SG.A-carry-PST **yesterday-ADJ** dance-POT-ABS
'I liked yesterday's dances.'

Adjectival adjuncts (31) are allowed with the potentialis in core cases (recall that these forms show more nominalized properties than those with non-core cases), and in the absence of person and number agreement prefixes. That is, these constructions show a high degree of nominalization.

The potentialis is compatible with postpostions, even when agreement prefixes are present, cf:

(32) *se* *s-e-ḳ^{w}e* [*gazet* *qe-**s**-š'efə-nə-m* ***paje***].
I 1SG.ABS-DYN-go newspaper DIR-1SG.A-buy-POT-OBL **for**
'I am going [to a shop] to buy a newspaper.'

In conclusion, the Adyghe potentialis occupies a high position in all finiteness hierarchies, since it demonstrates nearly all the properties, including distributional properties, shown by a verb in the language. However, it can also head nominalized constructions – in these cases it shows nominal morphological features, e.g. it may take adjectives and the possessive marker, and may appear without agreement markers. In this paper, I do not analyze such constructions, since they are employed only marginally in complementation and have peculiar semantic nuances, as is characteristic of deverbal nouns with a high degree of lexicalization.

The ability to take core cases and postpositions does not signal a high degree of nominalization. However, as shown in (25)–(29), the potentialis with the core cases is more nominalized than the potentialis with non-core cases.

3.2 Morphosyntax of complement clauses headed by the verbal stem

The verbal stem in complement clauses preserves the full range of verbal morphology, with the exception of the mood and dynamicity markers: see examples (33)–(35) with the agreement prefixes, (33) with the refactive aspectual suffix, (34) with the negation marker. As for tense markers, the verbal stem takes all tense markers that are used in independent clauses, cf. (33) with the past tense marker and (35) with the future marker.

(33) *se* [*w-jə-dese-xe-r* ***p***-*ŝə-**ž'ə-ʁe**-xe-ç'e*] *s-je-neg^{w}əje*.
I 2SG.PR-POSS-lesson-PL-ABS **2SG**-do-**RE-PST**-PL-INS 1SG.ABS-OBL-think
'I hope that you have already done your homework.'

(34) *g^{w}əš'əʔe* *pəte* *qə-w-e-s-e-tə* [*pçə* ***sə-mə**-wəsə-ç'e*].
word hard DIR-2SG-OBL-1SG.A-DYN-give lie **1SG.A-NEG**-invent-INS
'I give you my word of honour that I'm not lying.'

(35) [*a* *ǯ'anə-r* *pŝaŝe-m* *ə-š'efə-š't-ew*] *qə-s-š'-e-χ^{w}e*.
DEM dress-ABS girl-OBL 3SG.A-buy-FUT-ADV think-DIR-1SG.IO-LOC-DYN-become
'I think that the girl will buy that dress.'

The unacceptability of mood markers and the absence of the dynamicity marker (34) are among the properties that differentiate these constructions from the independent clause construction. Another property of this kind is the use of subordinate negation (34).

As for nominal features, the verbal stem can take the markers of adverbial (35) and instrumental cases (34); core cases are impossible in the construction under discussion. The possessive marker is also unacceptable.

All the syntactic properties of the complement constructions with the verbal stem follow the clausal and not the nominal pattern: the verbal stem assigns to its core arguments the same cases as those required in independent clauses (cf. absolutive and oblique in [35]), and the adjuncts are encoded as adverbs, not as adjectives. The verbal stem can occur with postpositions.

As shown in Section 3.1, core cases signal a higher degree of nominalization than the adverbial and the instrumental cases (and postpositions). Hence, it can be concluded that the verbal stem in complement clauses does not demonstrate nominal morphological properties.

3.3 Morphosyntax of complement clauses headed by the factive form

The factive form (traditionally termed the participle (cf. Rogava & Kerasheva 1966: 111) is formed with the prefix *zere-* and takes the case marker required by the complement-taking predicate:

(36) *ç'elejeʁaǯ'e-m* *je-ŝe* [*azemat* *ŝhanʁ^{w}əpče-r* ***zere**-x-jə-wətə-ʁe-**r***].
teacher-OBL OBL+DYN-know Azamat window-ABS **FCT**-LOC-3SG.A-break-PST-**ABS**
'The teacher knows that Azamat has broken the window.'

The factive form preserves verbal morphological features, including polypersonal agreement, aspect, dynamicity marking, valency-changing affixes, and verbal

tenses; cf. examples with the agreement markers (36)–(37), with the refactive aspectual suffix (38), with negation (37), with the dynamicity marker *-re-* (37) and with the markers of the past (36) and future (38) tense. Mood markers are unacceptable on the factive form. Another property differentiating the factive form from verbal forms that head independent clauses is the use of subordinate negation *mə-* (37) and dynamicity marker *-re-* (37).

(37) [*ə-ne-me* *deʁʷə* *zer-a-**mə**-ʎeʁʷə-**re**-m* *paje*]
3SG.PR-eye-OBL.PL good FCT-3PL.A-**NEG**-see-**DYN**-OBL for
a-š' *mašine* *ə-fə-r-ep*.
DEM-OBL car 3SG.A-drive-DYN-NEG
'He doesn't drive the car, because he has poor eyesight.'

(38) [*zere-ḳʷe-**ž'ə**-š'te-m-ç'e*] *q-a-r-jə-ʔʷa-ʁ*.
FCT-GO-**RE**-FUT-OBL-INS DIR-3PL-AUG-3SG-say-PST
'He told them he would leave. (lit. As for his departure, he told them about it).'
(Gerassimov & Lander 2008: 293)

The main nominal feature of the factive form is the presence of the nominalizer *zere-*. Like the potentialis form, this can take all four nominal case markers. However, it cannot take the possessive marker or the nominal plurality marker; adjectival modifiers are also unattested with the factive form.

The syntax of the complement clause headed by the factive form follows the clausal pattern. Both core arguments and adjuncts are encoded in the same way as in independent finite clauses. However, the factive form can take all four cases and may occur with postpostions, cf.:

(39) *a-r* *me-gʷəṣʷe* [*xə-m* *zere-ḳʷe-š'tə-m* **paje**].
DEM-ABS DYN-hope sea-OBL FCT-GO-FUT-OBL **for**
'He is glad because he's going to travel to the sea.'

Hence, the factive form shows mostly verbal properties, except for the possibility of taking core case markers and appearing in the context of postpositions.

3.4 Morphosyntax of the complement clauses: summary

The properties of the three forms discussed in sections 3.1–3.3 are summarized in Table 1 below.

Table 1: Morphosyntactic properties of the potentialis, the factive forms and the verbal stem with case markers in complement clauses in Adyghe

Nominal / verbal properties	Potentialis	Factive form	Verbal stem with case markers
Verbal properties			
mood	–	–	–
tense markers	+/–	+	+
aspect- and valency-changing markers	+	+	+
agreement	+/–	+	+
negation	non-finite[8] negation	non-finite negation	non-finite negation
dynamicity	–	+	–
core arguments encoded in the same way as in independent clauses	+	+	+
adjuncts marked as adverbs	+	+	+
Nominal properties			
nominalizer	+	+	–
case markers	+	+	+/–
nominal plurality markers	n/a	n/a	n/a
adnominal marking of the arguments	–	–	–
possessive markers	+/–	–	–
compatibility with postpositions	+	+	+

It can be concluded that the most frequently used complementation strategies in Adyghe do not demonstrate a high degree of nominalization and/or desententialization (cf. Lehmann 1988; Givón 1990). The constructions with the most nominalized properties involve the potentialis without agreement markers; however, even in these constructions the argument structure and the aspectual properties of the nominalized verb are mostly preserved.

In the next section I consider the three complementation strategies discussed here with regard to their ability to take case markers, and argue for the necessity of distinguishing more than three complementation strategies in Adyghe.

8 This only concerns the potentialis in complement clauses; in independent sentences the finite negation marker is used (23).

3.5 Case markers as complementizers

3.5.1 The distribution of case markers with the main complement constructions

As has been shown above, all three of the strategies under discussion take nominal case markers. Traditional works on Adyghe identify four case markers: ergative with the marker *-m*, absolutive with the marker *-r*, instrumental with the marker *-ç̣'e* and adverbial with the marker *-ew*. The ergative in Adyghe, besides its function of marking agent-like arguments of transitive verbs, is also used to mark indirect objects, including arguments introduced by the version prefixes (as shown in [4]), and adjuncts of time, location, etc. As shown in 1.1, the ergative introduces a large number of argument and adjunct roles, and there can be many noun phrases marked with the ergative in one and the same clause in Adyghe. Thus, Letuchiy (2012) argues for another term for this case marker, namely the *oblique*. As stated in section 2.1, the absolutive and the oblique case mark core arguments of the verb, as opposed to the remaining two cases, which are most often used to introduce adjuncts.

The instrumental case, as well as marking instrumental meaning proper, is used to mark the meaning of motion through an area (prolative), as well as cause, temporal interval, means of transport, stimulus of emotion etc. (cf. Serdobolskaya 2011). The adverbial case suffix is attached not only to nouns (in contexts of manner, as in *to jump like a kangaroo*), but is also used to mark adverbs (40), nouns in adverbial position, heads of relative clauses, secondary predicates (41) and adverbial clauses (42) etc. (cf. Arkadiev et al. 2009 for details).

(40) *psənç̣'e – psənç̣'-ew; daxe – dax-ew*
easy easy-ADV beautiful beautiful-ADV
'easy' – 'easily'; 'beautiful' – 'beautifully'
(Rogava & Kerasheva 1966: 334)

(41) *se ap-**ew** sə-ḳ^{w}a-ʁ.*
I first-**ADV** 1SG.ABS-go-PST
'I went first.'

(42) [*janəwe jə-ʁ^{w}ənеʁ^{w}ə-m a-dež' ḳ^{w}a-ʁ-**ew***] *ʔe ŝhalə-m*
POSS+old.woman POSS-neighbour-OBL 3PL.POSTP-to go-PST-**ADV** hand mill-OBL
ṗestapχe qə-š'-jə-haž'ə-š'tə-ʁ.
porridge DIR-LOC-3SG.A-grind-AUX-PST
'Having gone to the neighbours, the old woman ground the flour to make porridge.'
(Kerasheva 1995: 167)

Thus, the interpretation of this marker as a "case" in the reference grammars is rather misleading. However, in this article I retain the term "case" for lack of any other well-established term.

In complementation, all four cases can be attached to the head of the complement clause. However, the factive form in *zere-* and the potentialis can take all four case markers, while the verbal stem can only take[9] the instrumental and adverbial case markers (see Table 2):

(43) *[*azemat* *qa-ḳ^{w}e-**r**] *qə-s-š'-e-χ^{w}ə*.
Azamat DIR-go-**ABS** DIR-1SG.IO-LOC-DYN-become
Intended meaning: 'It seems to me that Azamat is coming.'

Such examples are judged ungrammatical, although nominal arguments of this verb with the same role of "theme" do take absolutive case marking:

(44) *se* *a-**r*** *qə-s-š'-e-χ^{w}ə*.
I DEM-**ABS** DIR-1SG.IO-LOC-DYN-become
'I think so.'

The same distribution is observed for all complement-taking predicates that take a complement clause headed by the verbal stem or a nominal argument with the absolutive/oblique case. It can be concluded that nominal and clausal arguments in Adyghe require different case marking. This phenomenon has been attested in some Mongolic (see Knjazev 2009 for Kalmyk) and Finno-Ugric languages (see Miteva 2009 for Komi-Zyrjan).

Table 2: Case marking of the dependent predicate in different complementation strategies

Complementation strategy	Absolutive in *-r*	Oblique in *-m*	Instrumentalis in *-ç'e*	Adverbial case in *-ew*
Factive form in *zere-*	+	+	+	+
Verbal stem	–	–	+	+
Potentialis in *-n*	+	+	+	+

9 It should be specified that in relative clauses, verbs in Adyghe can take all the case markers in reference to the participant in a situation: *ḳ^{w}a-ʁe-r* [go-PST-ABS] 'the one that has gone', *ḳ^{w}a-ʁe-m/-mč'e/-ew* [go-PST-OBL/INS/ADV] 'the one that has gone/with the one that has gone/as the one that has gone', *ḳ^{w}e-š'tə-r* [go-FUT-ABS] 'the one that will go' etc. (Rogava & Kerasheva 1966: 314–315). I do not consider these constructions here, since they function as relative clauses and not sentential complements.

The table shows that complements take non-core cases, that is, cases that are predominantly used to mark nominal non-arguments. However, in complemention non-core cases are used not as adverbial clause markers, but as complementizers. Dependent clauses with instrumental/adverbial case markers remain arguments of complement-taking predicates (see the discussion after example [16]).

3.5.2 Terminological remarks on the verbal stem with case markers

As shown in Section 3.2, the verbal stem in this construction preserves most of the morphosyntactic properties of verbal forms that head independent finite clauses. The only marker of subordination in these constructions is the case marker. In Adyghe reference grammars (Zekokh 1976; Kerasheva 1984 and others), the verbal stem with the adverbial case marker *-ew* is termed the "converb" (as opposed to the potentialis with the adverbial case, which is known as the "supine"), while the verbal stem with the instrumental case marker, as far as I know, receives no special designation (it is called a "non-finite formation in *-nčʼe*" in Kerasheva 1984). This form can be related to the frustrative (antiresultative) mood in Adyghe. The frustrative is formed with the marker *-čʼe*, which attaches to the verbal stem after the suffixes of tense, plural agreement and dynamicity; it is used in antiresultative contexts (i.e. with the meaning of 'cancelled result'):

(45) *a-r* ***šxe-čʼe*** *xe-χʷ-e-šʼt-ep*.
DEM-ABS **eat-INS** GROW-FUT-NEG
'Even if he eats, he'll never gain weight.'
(Kuznetsova 2009: 311)

It is unclear whether the marker seen in (45) and that seen in sentential complements should be considered to represent the same unit. From the point of view of morphosyntax, they do not show different properties: they occupy the same position in the verb form (position 5 after the verbal stem, as defined in Arkadiev et al. 2009), and they do not trigger the *e~a* alternation[10] in the previous syllable (this alternation differentiates the manner nominalization marker *-čʼe* from all three of the other markers with this shape: the case marker *-čʼe*, the frustrative marker *-čʼe*, and the sentential argument marker *-čʼe*). However, the semantics of (45) and of complement clauses with the instrumental do not seem to be related. The frustrative denotes a situation with a cancelled result, while the complement clause marker *-čʼe* marks eventive and irrealis complements. The second type of context can be linked functionally with the frustrative; see Plungian (2001) for

10 See Smeets (1984); Arkadiev et al. (2009).

the combination of irrealis and antiresultative meaning in markers of the pluperfect. However, the link between the meanings of event and irrealis/antiresultative is not so obvious. It could be speculated that the event and irrealis readings arise from different sources. The event reading may have developed from the instrumental case in *-ç'e* via the manner meaning of the instrumental (although in fact this meaning is rather marginal in Adyghe). A shift from manner towards event in the meaning of complementizers is attested in a number of languages, e.g. Slavic. However, the exact grammaticalization path of complements in *-ç'e* is unclear and needs further research. I label these forms "frustrative" in independent clauses, and "verbal stem with the instrumental case" in complement clauses (even in irrealis clauses); the gloss, however, is the same in all these contexts (INS).

I am not using the term "converb" for the verbal stem with the adverbial case, nor do I use "supine" for the potentialis with the adverbial case as in Kerasheva (1984). The rationale for this solution is that, as I show below, the potentialis with the adverbial case is more appropriately analyzed as a combination of two markers (the potentialis and the adverbial marker), which is fully compositional semantically.

3.5.3 Variation of case marking in complementation

The case marker found on the verb heading the complement clause can be determined by the complement-taking predicate, cf. the same marking for nominal and clausal arguments with the complement-taking predicate 'like':

(46) *ç'ale-m* *pŝaŝe-**r*** *ə-g^{w}* *r-j-e-hə*.
boy-OBL girl-**ABS** 3SG.PR-heart LOC-3SG.A-DYN-carry
'The boy likes the girl.'

(47) [*c^{w}əmpe-r* *qe-sə-wəʁ^{w}ejə-nə-**r***] *sə-g^{w}*
strawberry-ABS DIR-1SG.A-collect-POT-**ABS** 1SG.PR-heart
r-j-e-hə-r-ep.
LOC-3SG.A-carry-DYN-NEG
'I don't like to collect strawberries.'

Both nominal and clausal arguments take the absolutive case in (46) and (47), just as the same preposition is used in the English sentences *I'm afraid **of** him* and *I'm afraid **of** his doing nothing about the work*. For the sake of brevity, I will henceforth use the term *subcategorized case marking* for case marking that follows the pattern shown by nominal arguments.

However, complement clauses in Adyghe may instead take other case markers which are not determined by the complement-taking predicate (non-subcatego-

rized case marking): for example, the verb *əgʷ rjəhə-* 'like' in (48) takes the potentialis in the instrumental or adverbial case:

(48) a. [*cʷəmpe-r* *qe-sə-wəʁʷejə-n-**ç'e*** / *qe-sə-wəʁʷejə-n-**ew***]
strawberry-ABS DIR-1SG.A-collect-POT-**INS** DIR-1SG.A-collect-POT-**ADV**
sə-gʷ *r-j-e-hə*.
1SG.PR-heart LOC-3SG.A-DYN-carry
'I like collecting / to collect strawberries.'

Such variation is prohibited with nominal arguments:

b. **ç'ale-m* *pŝaŝe-**ç'e*** / *pŝaŝ-**ew*** *ə-gʷ* *r-j-e-hə*.
boy-OBL girl-**INS** girl-**ADV** 3SG.PR-heart LOC-3SG.A-DYN-carry
Intended meaning: 'The boy likes the girl.'

Let us consider another example of this variation. The verb *š'eš'əne-* 'fear' has the argument structure <ABS(experiencer) OBL(stimulus)>. Thus if the stimulus is an object, it is encoded by the oblique case:

(49) a. *se* *təʁʷəẑə-**m*** *sə-š'-e-š'əne*.
I wolf-**OBL** 1SG.ABS-LOC-DYN-fear
'I'm afraid of the wolf.'

Other case markers in this context are ungrammatical:

b. **se* *təʁʷəẑə-**ew*** / *-**ç'e*** / *-**ər*** *sə-š'-e-š'əne*.
I wolf-**ADV** /-**INS** / -**ABS** 1SG.ABS-LOC-DYN-fear
Intended meaning: 'I'm afraid of the wolf.'

The sentential complement, however, can be marked either with the oblique case, or with the instrumental or adverbial case:

(50) *se* [*č'eš'ə-m* *s-jə-zaqʷ-ew* *qe-s-ḳʷə.ha-nə-**m** /* *n-**ew** /* *n-**ç'e***]
I night-OBL 1SG-POSS-alone-ADV DIR-1SG.A-go-POT-**OBL** POT-**ADV** POT-**INS**
sə-š'-e-š'əne.
1SG.ABS-LOC-DYN-fear
'I'm afraid of walking around alone at night.'

(With this verb the instrumental and adverbial cases are regarded by native speakers as stylistically marked, but acceptable.)

The variation in complement case marking is limited by the following rules.

Unsubcategorized case marking only concerns non-core cases. The core cases are used in accordance with the same rules as nominal arguments: if the verb requires the oblique case, as here with the verb 'fear', it cannot assign absolutive case to its argument, whether this is nominal or sentential. The unsubcategorized non-core cases are acceptable with the sentential complement, as in (50). The

same goes for those verbs requiring the absolutive: e.g. the verb 'like' requires absolutive case on the stimulus, and it cannot assign oblique case to this argument, whether nominal or sentential; however, non-core cases are acceptable with the sentential complement.

Complement-taking predicates are found that cannot assign core cases to the stimulus, and require the instrumental or adverbial case, e.g.:

(51) *maše ja-dež'* [*ḳʷe-ž'ə-n-**ew***] *z-j-e-ʁe-hazərə.*
Masha 3PL.POSTP-to go-RE-POT-**ADV** REFL.ABS-3SG.A-DYN-CAUS-ready
'Masha is going to return home (lit. go back to hers).'

These predicates can only take complements with non-core cases. Some take only one, e.g. instrumental or adverbial case (such as *zəʁehazərə-* 'intend, be about to'), but there are also complement-taking predicates that may take either case (such as *gʷəʁe-* 'hope').

(52) *se* [*univjersitjetə-m* *sə-çe.fe-n-**ew*** / *-**ç'e***] *s-e-gʷəʁʷe.*
I University -OBL 1SG.ABS-enter-POT-**ADV** **INS** 1SG-DYN-hope
'I hope to enter the University.'

The restrictions formulated above can be summarized as follows. The use of the core cases with sentential complements conforms to the subcategorization of the matrix predicates; non-core cases can have non-subcategorized uses with sentential complements; nominal arguments do not allow this variation.

Variation in the case marking of complements is also restricted by complementation type: it is only observed with the potentialis and the verbal stem. As for the factive form, there is variation of oblique and instrumental with some emotive complement-taking predicates (for example, *jezeš'ə-* 'be fed up' and *əgʷ rjehə-* 'like', see the Appendix). This variation is more limited than that seen with the verbal stem and the potentialis, and I do not consider it in this paper.

Hence, the number of complementation strategies in Adyghe interacts with the possibility of case variation to give the following:

1a. Factive form in *zere-* with subcategorized case (oblique/absolutive/instrumental/adverbial)
1b. Factive form with non-subcategorized case: instrumental
2a. Verbal stem with the instrumental case
2b. Verbal stem with the adverbial case
3a. Potentialis in *-n* with one of the core cases (subcategorized marking)
3b. Potentialis in *-n* with the (non-)subcategorized instrumental
3c. Potentialis in *-n* with the (non-)subcategorized adverbial case

The question arises of how to characterize the semantic distinction between verbal forms which differ only in case marking: namely, between types 1a and 1b; 2a and 2b; 3a, 3b and 3c. This difference is rather subtle, and is not easily explained by native speakers. For Kabardian, where similar variation is attested, Kumakhov and Vamling (1998: 126) characterize the distribution of case markers in terms of interchangeability and vagueness of semantic difference. I will address this question in sections 4.2–4.3.

4 Semantics of complementation strategies

In this section I consider the functions of the factive from in *zere-*, the potentialis and the verbal stem with case markers.

4.1 Semantics of the factive form

4.1.1 Fact vs. proposition

The factive form is used in classical factive contexts, as described in theoretical works (Kiparsky and Kiparsky 1971; Arutjunova 1988; Zalizniak 1990); cf. Serdobolskaya (this volume) for the overview of tests that identify facts in complementation. This complementation strategy is used with the factive verb *ŝ́e-* 'know' (53) and the emotive factive verbs *jeχ^{w}epse-* 'envy' (54), *(fe)g^{w}əbžə-* 'be angry' and others (see Gerassimov & Lander 2008 for details).

(53) [*ç̌əg^{w}ə-r* ***zere***-*χ^{w}əraje-r*] *zeç̌'e-m-jə* *ja-ŝ́e*.
Earth-ABS **FCT**-round-ABS all-OBL-ADD 3PL.A+DYN-know
'Everybody knows that the Earth is round.'

(54) *pŝaŝe-r* *jeχ^{w}epse* [*ə-šəpχ^{w}* *ç̌'ale-m* *nah* *ə-g^{w}*
girl-ABS envy 3SG.PR-sister boy-OBL than 3SG.PR-heart
zere-*r-jə-hə-re-r* *jež'* *nah-rjə*].
FCT-LOC-3SG.A-carry-DYN-ABS INTF/REFL than-PRT
'The girl envies her sister because the boy likes her sister more than her.'

Non-factive mental verbs of opinion meaning 'think, believe' (*ləte-*, *š'eχ^{w}ə-* etc.) do not allow the factive form (55b).

(55) a. [*ç̌əg^{w}ə-r* *χ^{w}əraj-**ew***] *s-e-ləte*.
Earth-ABS round-**ADV** 1SG.ABS-DYN-count
'I think that the Earth is round.'

b. *[*ç'əgʷə-r* ***zere***-*χʷəraje-r*] *s-e-ɬəte*.
Earth-ABS **FCT**-round-ABS 1SG.ABS-DYN-count
'I think that the Earth is round.'

On the contrary, complements in the adverbial case, which do not have a factive reading, are preferred with these verbs (55a), and in most contexts are judged as unacceptable with *ŝe-* 'know' (56b).

(56) a. [*azemat* *qə-****zere***-*ḳʷe-š'tə-r*] *s-e-ŝe*.
Azamat DIR-**FCT**-go-FUT-ABS 1SG.A-DYN-know
'I know that Azamat will come.'
b. *[*azemat* *qe-ḳʷe-n-****ew***] *s-e-ŝe*.
Azamat DIR-go-POT-**ADV** 1SG.A-DYN-know
'I know that Azamat will come.'

When the verb 'know' is used in the non-factive meaning 'be certain', 'be sure' (57) or in presupposition-opaque contexts (58) (cf. Serdobolskaya, this volume for details), it can take the potentialis with non-core cases, cf.:

(57) [*azemat* *qe-ḳʷe-n-****ew***] *w-e-ŝ-a*?
Azamat DIR-go-POT-**ADV** 2SG.A-DYN-know-Q
'Are you sure that Azamat will come?'

(58) [*psewənəʁe* *w-jə-ʔe-n-****ç'e***] *s-ŝe-me*
health 2SG.A-LOC-be-POT-**INS** 1SG.ABS-know-COND
qə-p-fe-sə-mə-ŝe-n *š'ə-ʔ-ep*.
DIR-2SG.IO-BEN-1SG.A-NEG-do-POT LOC-be-NEG
'I wouldn't spare anything for your health (lit. There is nothing I wouldn't give, if I knew that you were healthy).'

Here the truth of the complement clause is not presupposed – in (57) it is questioned, and in (58) it is merely hypothesized by the speaker. In such contexts the factive form is not used. In (57) the dependent clause is propositional, and the adverbial case is used. The context of (58) suggests that the complement clause is in fact false (according to the native speakers' judgement), and the instrumental case is chosen. For the use of the adverbial case to encode propositions, and the instrumental to encode false propositions, see sections 4.2 and 4.3.

As well as encoding facts with factive verbs, the factive form can also be used with non-factive complement-taking predicates, e.g. 'love, like' and the verb of speech 'say':

(59) [*njepe* *weš'x* *q-je.š'xə-n-****ew***] *a-ʔʷa-ʁ*.
today rain DIR-rain-POT-**ADV** 3PL.A-say-PST
'They said it would be raining today.' (The sentence describes a radio forecast. The speaker does not know if the forecast will turn out true or false.)

(60) *s-jane* *je-s-ʔʷa-ʁ-ep* [*ocenke* *dej*
1SG.PR-POSS+mother OBL-1SG.A-say-PST-NEG mark bad
*qə-**zere**-s-hə-ʁe-r*].
DIR-**FCT**-1SG.A-carry-PST-ABS
'I haven't told mother that I got a bad mark.' (The speaker did receive a bad mark.)

Both examples contain complement clauses with the verb of speech, but in (59) it takes the potentialis with the adverbial case, while in (60) the factive form is used. The difference between the two examples corresponds to the difference between presupposition and assertion. In (59) the complement clause belongs to the assertion being made, while in (60) the matrix clause is asserted, and the complement is presupposed.

Complements with the factive form pass the diagnostic tests for facts. First, the negation test: for example, in (60) the negation of the complement-taking predicate does not affect the truth of the complement. In the judgment of native speakers, the truth of the complement clause is preserved in such cases, unlike in complement clauses with the verbal stem.

Second, the truth of the complement clause cannot be denied in the further context by the same speaker. If the speaker does not intend to present as true the situation described in the complement clause, the factive form cannot be employed, and other complementation strategies must be used instead, cf.:

(61) [*a-r* *qə-s-e-wa-ʁ-**ew***] *zeç'e-m-jə* *a-r-jə-ʔʷa-ʁ*.
DEM-ABS DIR-1SG.IO-OBL-hit-PST-**ADV** all-OBL-ADD 3PL.IO-OBL-3SG.A-say-PST
'He told everyone that he had hit me. {But this is not true.}'

Let us consider quasi-performative contexts. The factive form cannot be used in such cases: these contexts require the verbal stem with the adverbial case.

(62) a. *se* *azemat* *jane-jate-xe-m* *ja-s-ʔʷa-ʁ* [*azemat*
I Azamat mother-father-PL-OBL 3PL.AGR+OBL-1SG.A-say-PST Azamat
mə-š' ***zere**-š'-je-mə-ʒ'e-ž'ə-š'tə-r*].
DEM.PROX-OBL **FCT**-LOC-OBL-NEG-read-RE-FUT-ABS
'I told Azamat's parents that he would not study here.'
b. ...[*azemat* *mə-š'* *š'-je-mə-ʒ'e-ž'ə-š't-**ew***] *ja-s-ʔʷa-ʁ*.
Azamat DEM.PROX-OBL LOC-OBL-leave-RE-FUT-**ABS** 3PL.AGR+OBL-1SG.A-say-PST
'I told Azamat's parents that he would not study here.'

In (62a) the sentence is interpreted by native speakers as follows: the decision that Azamat will not study at this school has been taken before the moment described in the matrix clause, and the speaker (and perhaps the hearer) is aware of this. However, (62b) can be interpreted in such a way that the speaker is the director of the school, and the decision was taken at the moment of the speech event described in the matrix clause, and was in fact implemented in the statement "He

will not study here" itself. This quasi-performative context takes a proposition in the complement clause, and hence the adverbial case complement is used.

Therefore, the conclusion can be drawn that the factive form introduces facts, while forms using the adverbial case introduce propositions.

4.1.2 Fact vs. event

Let us consider minimal pairs with factive and eventive complements. In (63) what is being evaluated by the speaker is the fact that Aslan has come, while in (64) s/he is evaluating the emotions and physical sensations arising from the activity of walking.

(63) [*asʎan* *qə-**zere**-ḳ^{w}a-ʁe-r*] *deʁ^{w}ə*.
Aslan DIR-**FCT**-go-PST-ABS good
'It is good that Aslan has come.'
(Gerassimov & Lander 2008: 292)

(64) [*pč'əha.re* *qe-p-ḳ^{w}ə.ha-n-**ç'e***] *deʁ^{w}ə*.
in.the.evening DIR-2SG.A-go-POT-**INS** good
'It's nice walking in the evening.'

The factive form can be used in (63), and not in (64), because (64) introduces the eventive context (cf. the tests for distinguishing between propositions, facts, and events in Serdobolskaya, this volume). The same semantic opposition is observed in (65) and (66): in (65) the speaker describes as positive his feelings arising from the situation, while in (66) s/he is evaluating the fact of it being warm as positive. Hence, in (66) the factive form is used, in contrast to (65).

(65) [*c^{w}əmpe-r* *qe-sə-wəʁ^{w}ejə-n-**ç'e***] *sə-g^{w}* *r-j-e-hə*.
strawberry-ABS DIR-1SG.ABS-collect-POT-**INS** 1SG.PR-heart LOC-3SG.A-DYN-carry
'I like collecting strawberries. (I enjoy this activity.)'

(66) [*njepə* ***zere**-fabe-r*] *sə-g^{w}* *r-j-e-hə*.
today **FCT**-warm-ABS 1SG.PR-heart LOC-3SG.A-DYN-carry
'I like the fact that it is warm today.'

Therefore, the factive form denotes facts, while forms with the instrumental case denote events.

4.1.3 Other functions of the factive form

Other functions of the factive form are peripheral (cf. Gerassimov & Lander 2008). First, it is found in the context of topical irrealis complements, cf:

(67) [*marine* *bzəlfəʁe* *daḳʷə* ***zere***-*χʷə-š'tə-r*] *a-rə-ʁe-š'tə-n*.
Marina woman seamstress **FCT**-become-FUT-ABS DEM-PRED-PST-AUX-POT
'It is likely that Marina will become a seamstress.'
(Gerassimov & Lander 2008: 299)

According to Gerassimov and Lander (2008), the factive form is chosen in (67) because the complement clause constitutes the topic of the sentence, cf. Section 4.2.3 for the role of information structure in the choice of complementizer.

Second, the factive form can mark complement clauses with manner semantics, and the eventive complements of verbs of immediate perception:

(68) *se* *sə-gʷ* *r-j-e-hə* [*a-š'* *wered*
I 1SG.PR-heart LOC-3SG.A-DYN-carry DEM-OBL song
*qə-**zer**-jə-ʔʷe-re-r*].
DIR-**FCT**-3SG.A-say-DYN-ABS
'I like the way he sings.'

(69) [*fatime* *qə-**zere***-*ŝʷe-re-r*] *hasanə* *zə-leʁʷə-m* *ŝʷə* *ə-leʁʷə-ʁ*.
Fatima DIR-**FCT**-dance-DYN-ABS Khasan TEMP-see-OBL good 3SG.A-see-PST
'When Khasan saw Fatima dancing, he fell in love with her.'

These uses are explained by the diachronic origin of the factive form. Gerassimov and Lander (2008: 307–310) argue that it stems historically from the relative construction, where the instrumental argument is relativized. This analysis is supported by the fact that the prefix *ze-* functions as a relativizer in Adyghe ("participle" in reference grammars), while *re-/rə-* is a valency-increasing prefix which introduces an instrumental argument. See the contrast between (70), without the instrumental argument, and (71), where the instrumental argument is introduced by *rə-*.

(70) *mašine-r* *psənč'-ew* ***ma-ḳʷe***.
car-ABS fast-ADV **DYN-go**
'The car goes fast.'
(Rogava & Kerasheva 1966: 334)

(71) *kʷə-xe-r* *ʁʷəgʷə-m* ***r-e-ḳʷe-x***.
cart-PL-ABS road-OBL **INS-DYN-go-PL**
'The carts are going down the road.'
(Jakovlev & Ashxamaf 1941: 66, cit. after Gerassimov & Lander 2008: 308)

Gerassimov and Lander (2008) suggest that the manner interpretation (68) of the factive form arises from this diachronic source. I suggest that the manner interpretation, in its turn, gives rise to the use of the factive form in eventive contexts with immediate perception verbs (69)[11].

4.1.4 Some conclusions

Thus, the main function of the factive form in Adyghe is to denote facts; it can also introduce topical irrealis propositions, manner complements and events with immediate perception verbs. It is therefore used with the factive verbs *ŝe-* 'know', *š'əʁʷəpše-* 'forget', *jeχʷepse-* 'envy', *fegʷəbžə-* 'be angry', *gʷəŝʷe-* 'rejoice', and with non-factives that allow factive complements: emotive verbs (as predicates with the semantics 'love, like' and *jezeš'ə-* 'fed up'), the verbs of perception *ləʁʷə-* 'see' and *zexexə-* 'hear', mental verbs, and the verb *ʔʷe-* 'say'. It is noteworthy that facts are only encoded by means of the factive form.

As I show below, the distribution of the potentialis and the verbal stem in complementation cannot be explained on the basis of the opposition of semantic parameters discussed in (Serdobolskaya, this volume): propositions vs. events, irrealis complement clauses vs. propositions etc. The difference between the potentialis and the verbal stem in complement clauses belongs to the domain of temporal reference of the dependent clause. Meanwhile, events and propositions are differentiated by means of the case markers occurring on the complement.

In the following sections I consider only non-factive contexts, since facts are encoded by the factive form only.

4.2 Opposition of adverbial and instrumental case in complement clauses

The verbal stem and the potentialis in *-n* can take both adverbial and instrumental markers with one and the same complement-taking predicate. The distribution of the case markers seems arbitrary at first sight, since there are many matrix predicates that allow both case markers without any apparent difference in meaning. Analyzing the same phenomenon in Kabardian, a language belonging to the same Circassian subgroup of the Northwest Caucasian language family, Kumakhov and Vamling (1998: 126) claim that the instrumental and adverbial

11 The shift from manner to event in the meaning of complementizers is observed in Slavic languages, cf. Arutjunova (1988) for Russian.

case markers are very close in meaning and interchangeable in complement clauses. Native speakers of Adyghe often see no difference in meaning between the two cases. However, there are contexts where one variant is preferred by all speakers. Moreover, not all matrix predicates allow variation in case marking. I suggest that the distribution of the cases in complement constructions is based on the opposition between events and propositions.

4.2.1 Potentialis / verbal stem with the adverbial case

As shown in Section 4.1.1, in many contexts the factive form is semantically opposed to the forms with the adverbial case (the verbal stem and the potentialis with the adverbial case), see examples (55)–(56), (59)–(60) and (62). In (61) the complement clause with the adverbial case is used to denote a proposition that is negated in the following context (unlike with factive complements, which cannot be negated in the following context). In (62b) it introduces a proposition in a quasi-performative context. Hence, forms with the adverbial case meet the criteria for propositions discussed in (Serdobolskaya, this volume).

There is additional evidence for the claim that forms with the adverbial case (both the verbal stem and the potentialis) denote propositions. The potentialis with the adverbial case introduces purpose clauses (72) and complements of speech causation verbs (73), see Asher (1993) for the analysis of complements of such complement-taking predicates as propositions.

(72) *se daḳʷe-m šeç' qə-fe-s-hə-ʁ* [*ʒ'ane qə-s-f-jə-də-**n-ew***].
I tailor-OBL cloth DIR-BEN-1SG.A-carry-PST dress DIR-1SG.IO-BEN-3SG.A-SEW-**POT-ADV**
'I brought the cloth to the tailor in order for him to sew a dress.'

(73) *se je-s-ʔʷa-ʁ azemat* [*urokə-m ç'e-ç'ə-ž'ə-**n-ew***].
I OBL-1SG.A-say-PST Azamat lesson-OBL LOC-leave-RE-**POT-ADV**
'I told Azamat to leave the class.'

Note that the forms with the instrumental and with the adverbial case are differentiated prosodically. Complements with the adverbial case usually host the main accent – a property that characterizes propositions in a number of languages (Yanko, p.c.; cf. Yanko 2001: 242–245 for Russian). Meanwhile, cross-linguistically events may or may not take the main accent, and Adyghe complements with the instrumental case show the same variation.

Unlike eventive complements, complements with the adverbial case can host negation:

(74) *a-r* *me-š'əne* [*jə-ç'elejeʒ'aḳʷe-xe-r* *univjersitjetə-m*
DEM-ABS DYN-fear POSS-pupil-PL-ABS University-OBL
*ç'e-mə-haŝʷə-n-x-**ew*** / ***ʔ-n-xe-ç'e***].
LOC-NEG-enter-POT-PL-**ADV** **POT-PL-INS**
'He fears that his pupils will not enter the University.'

The complement contains the negation prefix *mə-*, and native speakers prefer the variant with the adverbial marker in this context.

On the basis of these arguments, I claim that complements with the adverbial case in Adyghe denote propositions.

However, there are contexts where the function of forms with the adverbial case is less clear. The adverbial case marker is found with evaluative complement-taking predicates and with complement-taking predicates with the meaning 'love, like':

(75) [*we* *çəŝhe.məŝhe-xe-r* *p-šxə-n-**ew***] *deʁʷə.*
you fruit-PL-ABS 2SG.A-eat-POT-**ADV** good
'It is good for you (for your health) to eat fruit.'

(76) [*pč'edəž'ə-m* *ž'-ew* *sə-qe-teʒ'ə-n-**ew***] *sə-gʷ*
morning-OBL early-ADV 1SG.ABS-DIR-get.up-POT-**ADV** 1SG.PR-heart
r-j-e-hə.
LOC-3SG.A-DYN-carry
'I like getting up/to get up early in the morning.'

With these complement-taking predicates both the adverbial and the instrumental are possible. It has been claimed for many languages that these complement-taking predicates take factive (e.g. *I like that he protects me*; *It is good that he protects you*) or eventive complements (e.g. *I like dancing; Dancing is good for your health*). However, I assume that they can also take propositions. The evaluative predicates can be used when the complement denotes a situation imagined by the speaker, cf. Russian:

(77) *Xorošo,* ***čto,*** ***mož-et*** ***by-t'***, *et-o* *by-l* *l'edan-oj*
good **COMP** **can-PRS.3SG** **be-INF** this-N.SG.NOM be-PST(M.SG) ice-M.SG.NOM
meteorit.
meteorite
'It is good that it was possibly an ice meteorite.'

(78) *A* *vam* *po-nravi-l-o-s'* *by*, ***jesli*** *by* *vaš* *muž*
and you.DAT PFV-like-PST-N.SG-REFL SBJV **if** SBJV your husband
po-šel *v* *sportzal* *s* *ženščin-oj?*
PFV-go.PST(M.SG) in gymnasium with woman-INS
'Would you like it if your husband went to a gymnasium with a woman?'
(Russian National Corpus www.ruscorpora.ru)

The imagined situation in the complements with *čto* 'that' and *esli* 'if' is evaluated as positive. In both examples the complement is not eventive, since what is being evaluated is the pure fact (and not the feelings of the speaker). Moreover, the complements in both examples can host negation:

Russian

(77') *Xorošo,* ***čto*** *et-o* ***ne*** *l'edan-oj* *meteorit.*
good **COMP** this-N.SG.NOM **NEG** ice-M.SG.NOM meteorite
'It is good that it is not an ice meteorite.'

(78') *A* *vam* *po-nravi-l-o-s'* *by,* ***jesli*** *by* *vaš* *muž* ***ne***
and you.DAT PFV-like-PST-N.SG-REFL SBJV **if** SBJV your husband **NEG**
po-šel *v* *sportzal* *s* *ženščin-oj?*
PFV-go.PST(M.SG) in gymnasium with woman-INS
'Would you like it if your husband did not go to a gymnasium with a woman?'

These are not factive, since the complement can be negated in the following context by the same speaker:

(77'') *Xorošo,* ***čto,*** ***mož-et*** ***by-t',*** *et-o* *by-l*
good **COMP** **can-PRS.3SG** **be-INF** this-N.SG.NOM be-PST(M.SG)
ledan-oj *meteorit,* *xot'a* *ja* *v* *et-o* *ne* *ver'-u.*
ice-M.SG.NOM meteorite although I in this-N.SG.NOM NEG believe-PRS.1SG
'It is good that it was possibly an ice meteorite, although I do not believe it.'

(78'') M*ne* *ne* *po-nrav-it-s'a,* ***jesli,*** *(****mož-et*** ***by-t',****)* *moj* *muž*
I.DAT NEG PFV-like-PRS.3SG-REFL **if** **can-PRS.3SG** **be-INF** my husband
po-jd-et *v* *sportzal* *s* *ženščin-oj.* *Vr'ad li* *on* *tak*
PFV-go-PRS.3SG in gymnasium with woman-INS hardly he so
s-delaj-et.
PFV-do-PRS.3SG
'I will not like it if my husband goes to a gymnasium with a woman. He would hardly do such a thing.'

They can both contain an epistemic expression, which is a diagnostic for propositions (Boye 2012), see (77'') and (78'').

Therefore, these complements are to be analyzed as propositional. See (Letuchiy 2014) on the analysis of Russian if-complements with emotive complement-taking predicates as non-factives. It seems that similar contexts are observed in English constructions where the complement of the evaluative predicate is introduced by the conjunction *if*:

(79) *John might hate it* ***if*** *he won.*
(Quer 1999: 242)

In Spanish and Catalan, similar constructions are observed with the emotive verbs 'please', 'annoy', 'hate' and others:

Catalan

(80) *M'agrada* *molt* *pro* [***si*** *fas* *pastissos*].
me-please.3SG a-lot **if** make.2SG cakes
'I like it a lot if you make cakes.'
(Quer 1999: 242)

Quer shows that these constructions are often non-factive, even if they are attested with factive emotive verbs (Quer 1999: 252).

As the conjunction in question is believed to introduce conditional, i.e. adjunct clauses, examples similar to (79) and (80) are not often involved in studies of complementation (see Quer 1999 and 2008 for argumentation against the adjunct analysis), which is why it is believed that the complement-taking predicates under discussion do not take propositional complements. However, I claim that evaluative predicates in Adyghe do take propositional complements, as examplified in (75), and that Adyghe verbs with the meaning 'love, like' count positive evaluation among their meanings, which is why they may take propositional complements too (76).

4.2.2 Potentialis / verbal stem with the instrumental case

The instrumental case can be used in the following contexts. Firstly, it is found in the irrealis (81) or in false complements[12] with mental verbs (82).

(81) [*a* *ǯ'anə-r* *ə-š'efə-n-**ç'e***] *w-je-neg^{w}əj-a*?
DEM dress-ABS 3SG.A-buy-POT-**INS** 2SG.ABS-OBL+DYN-suppose-Q
'Do you think she'll buy this dress? (I doubt it.)'

(82) [a-*š'* *j-e-ṭə-**ç'e***] *j-e-ləte*.
DEM-OBL 3SG.A-DYN-dig-**INS** 3SG.A-DYN-think
a. 'He thinks he's digging. (He doesn't understand how poorly he is working.)'
b. 'He pretends to dig. (He 'potters': he pretends he's working, but he doesn't dig effectively.)'

In (81), the situation in the complement clause is interpreted as irrealis (as defined in Serdobolskaya, this volume): the speaker strongly doubts its validity. As for

12 This meaning was originally identified by Julia Kuznetsova.

(82), native speakers suggest two interpretations of this example, cf. the translations; in both cases the complement clause has the truth value 'false'.

The 'false' interpretation of the instrumental contrasts with the neutral truth value interpretation of the verbal stem with the adverbial case, cf.:

(83) a. [*s-šə-nahə.ç'e* *čəja-ʁe-**ç'e***] *me-g^{w}əʁe*.
1SG.PR-brother-younger sleep-PST-**INS** DYN-hope
'She thinks that my younger brother has fallen asleep. {In fact, he's still watching TV}.'
b. [*s-šə-nahə.ç'e* *čəja-ʁ-**ew***] *s-e-g^{w}əʁe*.
1SG.PR-brother-younger sleep-PST-**ADV** 1SG.ABS-DYN-hope
'I think that my younger brother has fallen asleep. (In reality, he may or may not have.)'

The irrealis reading is largely found with mental verbs of opinion (see also [58] with the verb 'know') and with singular verbs of causation and potential situation. For example, the verb *g^{w}əš'ʔe jetə-* 'promise, lit. give a word' most often takes the potentialis with the adverbial case; the instrumental is used if the speaker presumes or suspects that the promised situation will not be accomplished. Note that the irrealis reading is possible with both the potentialis (81) and the verbal stem (82).

Another type of context is that of events (see Serdobolskaya, Motlokhov 2009 for details), as demonstrated in (63)–(66) above, where the forms with instrumental marking contrast with the factive form. Consider some more examples with emotive verbs:

(84) a. *a-r* [*doske-m* *dež'* *qə-de-ç'ə-n-**ç'e***] *š'əne-zepətə-š'tə-ʁ*.
DEM-ABS blackboard-OBL to DIR-LOC-go-POT-**INS** fear-always-AUX-PST
'{Teachers often called him up, but} He was always afraid to go to the blackboard.'
b. *a-r* *me-š'əne* [*jə-ç'eleježʼaḳ^{w}e-xe-r* *univjersitjetə-m* *ç'e-mə-haŝ^{w}ə-n-x-**ew***].
DEM-ABS DYN-fear POSS-pupil-PL-ABS University-OBL LOC-NEG-enter-POT-PL-**ADV**
'He is worried about his pupils not getting into university.' (If this happens it will harm his pedagogical reputation.)

(85) a. [*we* *çəŝhe.məŝhe-xe-r* *p-šxə-n-**ew***] *deʁ^{w}ə*.
you fruit-PL-ABS 2SG.A-eat-POT-**ADV** good
'It is good for you (for your health) to eat fruit.'
b. [*sə-qə-b-de-g^{w}əš'ʔe-n-**ç'e***] *deʁ^{w}ə*.
1SG.ABS-DIR-2SG.IO-COM-talk-POT-**INS** good
'It is pleasant to talk to you.'

(86) a. [*cʷəmpe-r* *qe-sə-wəʁʷejə-n-**ew***] *sə-gʷ* *r-j-ə-hə-r-ep*
strawberry-ABS DIR-1SG.A-collect-POT-**ABS** 1SG.PR-heart LOC-3SG.A-DYN-carry-DYN-NEG
aw *varenje* *s-ʂə-n-ew* *sə-faj.*
but jam 1SG.A-do-POT-ADV 1SG.A-want
'{The girl has been collecting strawberries for hours. – Do you like collecting berries that much?} – Oh no, I don't like collecting strawberries, but I want to make jam.'

b. [*cʷəmpe-r* *qe-sə-wəʁʷejə-n-**ç'e***] *sə-gʷ* *r-j-e-hə.*
strawberry-ABS DIR-1SG.ABS-collect-POT-**INS** 1SG.PR-heart LOC-3SG.A-DYN-carry
'I like collecting strawberries. (I enjoy this activity.)'

The examples with the instrumental case are used in contexts where the complement-taking predicates in question denote ungrounded emotion without any mental processing: the experiencer in (84a) fears the situation of going to the blackboard at school independently of the possible harm that s/he may undergo as a result. In (84b), however, the experiencer fears the situation because of its possible consequences. This sentence can be paraphrased as 'the experiencer fears the consequences of the dependent situation, considering them as unwanted' (a similar characterization is given in Zalizniak 1992 for the meaning of the Russian verb *bojat'sa* 'fear'; the same difference is observed in the English examples *I'm afraid to go to the forest* vs. *He's afraid of receiving negative feedback*).

The predicate *deʁʷə* 'good' and the verb *əgʷ rjəhə* 'like' can take complements with both the adverbial and the instrumental case. With *deʁʷə* the instrumental is used if the predicate denotes emotion or physical pleasure; the adverbial case is compatible with the meaning of positive evaluation (from the point of view of morality, health, rationality or the like). The same distinction is observed with *əgʷ rjəhə*: the adverbial case is used in contexts where the complement-taking predicate denotes positive emotion based on evaluation (86a), the instrumental case being judged unnatural in these contexts, whereas the instrumental is preferred if the emotion is a positive feeling 'proper', as in (86b).

A similar semantic distinction is observed with most evaluative predicates: when denoting evaluation as such they take the adverbial case, whereas the instrumental case is used if the matrix predicate denotes a physical or emotional feeling. The difference observed for the verb *əgʷ rjəhə* 'like' is also attested with its close synonyms *jəç'as* and *ʂʷə ləʁʷə*.

Serdobolskaya and Motlokhov (2009) explain the observed difference in meaning on the basis of the distinction between events and propositions. Complements with the instrumental case introduce events, and those with the adverbial case introduce propositions. This explanation captures the semantic opposition observed in (84)–(86). In (85b) the aspects of the situation, i.e. the event itself, are evaluated as 'good', or giving pleasure acquired by the experiencer. By contrast, in (85a) the situation is mentally processed, and evaluated as 'good'. The role played by mental processing makes this situation a proposition. On

propositions with evaluative predicates and verbs of emotion see the discussion in Section 3.2.1.

This claim provides an explanation for the distribution of the instrumental and adverbial cases in the context of other matrix predicates. The instrumental case is generally unacceptable with matrix predicates that do not take eventive complements, i.e. verbs of causation, speech causation and potential situation (*pələ-* 'try', *(f)jeʁeʂe-* 'teach', *zəʁehazərə-* 'intend' etc.). On the other hand, with the predicate *tχaʁʷe* 'pleasant, delightful', which cannot denote evaluation proper, the instrumental case is the only possible non-core case:

(87) [*weš'x* *q-je.š'x-ew* *qe-p-ḳʷə.ha-n-**ç'e**/*-**ew***] *tχaʁʷe*.
rain DIR-rain-ADV DIR-2SG.A-go-POT-**INS**/*-**ADV** pleasant
'It is pleasant to walk in the rain.'

Most emotive and mental verbs can take either the instrumental or the adverbial case, depending on the semantic type of the dependent clause. The emotive verbs that allow for such variation are *š'əne-* 'fear', *faj* 'want', *χʷepse-* 'dream', *jezeš'ə-* 'pester' and *š'egʷəʁe* 'hope'. The verb *gʷəʂʷe-* 'rejoice' can take facts or propositions, and does not allow the instrumental case. The other emotive verbs that we have examined are factive (*jeχʷepse-* 'envy', *(fe)gʷəbžə-* 'be angry' and others), and only allow the factive form (54).

Mental verbs show the same semantic effect that has been observed with emotive verbs. Adyghe possesses a number of verbs meaning 'think, believe, suppose' (*ʂʷeʂə-, ləte-, jenegwəje-, š'eχwə-*). All these verbs take both instrumental case and adverbial case complements, cf.:

(88) a. *s-j-e-negʷəje* [*ɬʷə* *qe-s-hə-n-**ew***].
1SG.ABS-OBL-DYN-think two DIR-1SG.A-get-POT-**ADV**
'I think I'll get a bad mark (a "two").'
b. [*ç'ale-r* *ʂʷe-ḳʷedə-n-**ç'e***] *j-e-negʷəje*.
boy-ABS MAL-loose-POT-**INS** OBL-DYN-think
'She fears that she will lose this boy.'
c. [*məjeqʷape* *š'ə-psewə-**ç'e***] *s-j-e-negʷəje*.
Majkop LOC-live-**INS** 1SG.ABS-OBL-DYN-think
'I hope that he lives in Majkop.'

The adverbial case is used with the principal meaning of these verbs, i.e. the meaning that does not presuppose any emotional commitment on the part of the experiencer (88a). In these sentences, the dependent clause denotes a proposition. The instrumental case gives the dependent clause an eventive reading (or the falsity reading, as in [82] and [83a]). Mental verbs of opinion do not combine with eventive complements, cf. Russian **Ja polagaj-u, **kak** on ne prid'-ot* [I think-PRS.1SG how he NEG come-FUT.3SG] 'I think that he will not come' and English

**I think John's being late* in Rosenbaum (1967: 29). In Adyghe many mental verbs allow this type of complement, but with a particular semantic shift whereby the complement-taking predicate takes on a semantic component of emotional commitment, as in (88b) and (88c). The actual type of emotion – positive or negative – is not specified by the mental verb itself, and is inferred from the pragmatic context. This explains the translations given for (88b) and (88c): native speakers translate these examples with emotive verbs. Such a semantic shift is not observed in Russian or English mental verbs, which explains their inability to take eventive complements.

Note that the event reading (as well as the irrealis reading) is possible with both the potentialis (88b) and the verbal stem (88c). Meanwhile, generic events are encoded by the same means as events proper (86b), or by conditional and temporal subordinators, cf. Section 2.2. The generic event meaning is only attested with the potentialis.

Therefore, I conclude that the distribution of the instrumental and the adverbial in complement constructions can be accounted for in terms of the distinction between event and proposition. Complements with the instrumental case can also have the 'false' or 'irrealis' interpretation. The instrumental and adverbial case markers have thus become grammaticalized in Adyghe in the function of complementizers.

It should be specified that the use of the adverbial and the instrumental cannot be straightforwardly explained as the use of the factive form can. First, the observed semantic difference is very subtle with some complement-taking predicates and can barely be clarified even in the wider context. Second, some deviations are found from the pattern observed above. Some complement-taking predicates do not encode events by means of the instrumental: these are verbs of immediate perception, the verb 'forget' and phasal verbs.

Complement-taking predicates of immediate perception take the factive form (cf. 3.1.3) or the verbal noun in *-ç'e* (12) in eventive contexts. To encode propositions, the adverbial case is employed[13]:

(89) [*ə-š* *je-ŝʷ-**ew***] *s-e-ɬeʁʷə.*
3SG.PR-brother 3SG.IO-drink.ANTIP-**ADV** 1SG.A-DYN-see
'I see that his brother drinks [alcohol] / is a drinker.'

13 The meaning of the perception verb shifts towards *cognitive perception* (conclusion made from perceived events: e.g. the speaker has noticed a large number of empty bottles and general disorder in his brother's apartment, and has drawn the conclusion that his brother drinks heavily).

Forms with the instrumental case are not allowed with complement-taking predicates of perception. Likewise, phasal verbs (90) do not allow the potentialis with the instrumental case:

(90) *se* [*sə-če-n-**ew***] *(*sə-če-n-**ç'e**)* *s-je-ž'a-ʁ*.
I 1SG.ABS-run-POT-**ADV** 1SG.ABS-run-POT-INS 1SG.ABS-OBL-begin-PST
'I started running.'

In these contexts we see a deviation from the semantic distribution observed above. Thus, in general, the distribution of the markers is explained in terms of the opposition of propositions vs. events, but it appears that their grammaticalization in this function is still in progress, or has stopped at this point.

The question remains as to the function of the core cases with the potentialis. This question is addressed in Section 4.3. Another issue is the distribution of the verbal stem vs. the potentialis in complementation, which is addressed in Section 4.4.

4.2.3 Influence of information structure on the choice of complementation strategy

Example (67) shows that the choice of complementation strategy in Adyghe is also influenced by the information structure of the sentence. In (67), the complement clause is the topic of the sentence, and it belongs to the *irrealis complement* type. Irrealis complements are usually encoded by the verbal form with the instrumental case. However, in (67) the factive form is used. Hence, the topical position of the complement clause in the information structure is more important than its semantic type.

The information structure is also relevant for the choice of complementation strategy in the context of focused non-factive complements. If the complement is focused, or an element inside the complement is focused, the form with the adverbial case is preferable:

(91) *we* *nah* *w-jə-ç'ase-r* [*təʁe-m* *z-je-b-ʁewə-n-**ew***
you more 2SG.PR-POSS-favourite-ABS sun-OBL REFL.ABS-OBL-2SG.A-tan-POT-**ADV**
(-**ç'e**)*] *ar-a?* – *hawəmjə* [*zə-b-ʁe-psḳə-n-**ew***] *ar-a?*
(**INS**) DEM:PRED-Q or REFL.ABS-2SG.A-CAUS-bathe-POT-**ADV** DEM:PRED-Q
'Do you prefer tanning? Or bathing?'

The instrumental case is unacceptable here. By contrast, the instrumental case can be used if the complement-taking predicate constitutes the question focus:

(92) *w-e-š'əne* [*w-jə-zaq^{w}-ew* *wə-k̩^{w}e-n-**ç'e***] *ar-a*?
2SG.ABS-DYN-fear 2SG.PR-POSS-alone-ADV 2SG.ABS-go-POT-**INS** DEM:PRED-Q
'You're afraid of going alone, aren't you?' (Children are teasing a little girl.)

The position of the copula with the question marker is the same in both examples: it is placed after the dependent verb. However, in (92) it semantically modifies the complement-taking predicate, which is why the instrumental is possible.

Examples (91) and (92) contain a special focus construction described in (Sumbatova 2009); however, in Adyghe this construction is not obligatory in case of narrow focus, as shown in (93). Although the special focus construction is not used, the adverbial case is also preferable, because the complement is focused:

(93) *č'eš'nəq^{w}e-m* *we* *š'e-r* *te* *q-jə-p-xə-š't-a*? –
midnight-OBL you milk-ABS where DIR-LOC-2SG.A-take-FUT-Q
[*t-jə-ʁ^{w}ənеʁ^{w}* *ŝ^{w}əzə-m* *jə-ʔe-n-**ew***] *s-e-g^{w}əʁe*.
1PL.PR-POSS-neighbour married.woman-OBL LOC-be-POT-**ADV** 1SG.ABS-DYN-hope
'Where will you find milk at night? – I hope that the neighbour woman has some.'

Hence, the adverbial case is preferred in contexts where the complement constitutes the focus of the whole sentence or contains a focused element. Complements with the instrumental case are preferably not focused. This is in agreement with the fact that complements with the instrumental are used in topical constructions with reduplication:

(94) *je-ǯ'e-n-**ç'e*** / *-š't-**ç'e*** *j-e-ǯ'e* *awpar-jə*
OBL-read-POT-**INS** FUT-**INS** 3SG.ABS-DYN-read but
qə-g^{w}ərəʔ^{w}e-r-ep.
NPI-ADD DIR-understand-DYN-NEG
'He does read (lit. as for reading, he reads), but he doesn't understand anything.'

The adverbial case is unacceptable in this context.

Note that topicality in complement clauses can also be marked with the instrumental attached to the oblique (Serdobolskaya 2011).

Thus, non-core cases in complementation can have the function of signalling certain types of information structure: the instrumental can mark topical complements, while the adverbial marks focused complements.

4.3 Distribution of core and non-core cases with the potentialis

As argued in Section 3.2, the core cases are not available with the verbal stem in complement clauses. Meanwhile, variation of cases with the factive form is

attested with four complement-taking predicates only, and is not considered in this paper. Thus competition between the core and non-core cases occurs mostly with the potentialis. Syntactically, the two types of construction differ as follows. Complement clauses with the potentialis in the core cases are nominalized to a higher degree: they allow the omission of the prefixal agreement markers, and can take possessive prefixes and adjectival modifiers. Moreover, these constructions apparently form syntactic islands in terms of (Ross 1967): unlike the constructions with non-core cases, they do not allow relativization of the arguments of the embedded verb, cf.:

(95) a. [*mašin-ew* *p-fə-n-**ç'e/-ew***] *qə-p-fe-qjənə-r*
car-ADV 2SG.A-drive-POT-**INS**/**ADV** DIR-2SG.IO-BEN-hard-ABS
qe-t-š'efə-ʁ-ep.
DIR-1PL.A-buy-PST-NEG

b. *[*mašin-ew* *p-fə-n-**ər***] *qə-p-fe-qjənə-r*
car-ADV 2SG.A-drive-POT-**ABS** DIR-2SG.IO-BEN-hard-ABS
qe-t-š'efə-ʁ-ep.
DIR-1PL.A-buy-PST-NEG
'We haven't bought the car that is difficult to drive.' (lit. '...bought the car that it is difficult to drive')

In (95) the single argument of the embedded verb is the target of relativization. If the embedded verb is encoded by the potentialis with the instrumental/adverbial, relativization is allowed; however, if it occurs in the absolutive case, relativization is not allowed (95b).

I claim that this empirical difference is due to the nominalized properties of the potentialis with core cases.

The use of the potentialis with the core cases is similar to the use of (de)verbal nominals cross-linguistically. For example, the verb *pələ-* 'try' can have two meanings, 'be engaged in an activity' and 'try, attempt'. The first meaning is observed with nominal and clausal arguments:

(96) a. *sportə-**m*** *pələ-n*
sports-**OBL** be.engaged.in-POT
'to be engaged in sports'
(Txarkaxo 1991: 198)

b. [*adəga-bze-ç'e* *s-je-ʒ'e-nə-**m***] *sə-pəl.*
Adyghe-language-INS 1SG.ABS-OBL-read-POT-**OBL** 1SG.ABS-be.engaged.in
'I'm engaged in reading in Adyghe.'

Nominal arguments are marked with the oblique case (other cases are unacceptable). Clausal arguments with this meaning are also marked with the oblique. However, the complement with the adverbial case is acceptable if the complement-taking predicate is used in the second meaning:

(97) [*adəga-bze-ç'e* *s-je-ǯ'e-n-**ew***] *sə-pəλ*.
Adyghe-language-INS 1SG.ABS-OBL-read-POT-**ADV** 1SG.ABS-try
'I try to read in Adyghe.'

In this meaning, the verb in question does not take nominal arguments.

The same distinction is present in the pair of verbs *zəfeʁehazerə-* 'prepare, pack one's things' and *zəʁehazerə-* '1. intend, be going to, 2. prepare, pack one's things'. The first verb takes arguments in oblique, both clausal and nominal (the prefix *fe-* in Adyghe is a version prefix that introduces an argument in the oblique case). The second verb only takes clausal arguments, and only in the adverbial case (98c).

(98) a. *lekcije-**m*** *zə-fe-ʁe-hazərə-n*
lecture-**OBL** REFL.ABS-BEN-CAUS-ready-POT
'to prepare oneself for the lecture'
(Txarkaxo 1960: 160)

b. *azemat* [*wered* *q-ə-ʔʷe-nə-**m***] *zə-f-j-e-ʁe-hazərə*.
Azamat song DIR-3SG.A-say-POT-**OBL** REFL.ABS-BEN-3SG.A-DYN-CAUS-ready
'Azamat is preparing to sing.' (He may not yet have gone onstage, but he is getting out the microphone, guitar etc.)

c. *azemat* [*wered* *q-ə-ʔʷe-n-**ew***] *ze-ʁ-e-hazərə*.
Azamat song DIR-3SG.A-say-POT-**ADV** REFL.ABS-CAUS-DYN-ready
'Azamat intends to sing.' (He may already be on the stage and will be singing in a few seconds.)

The examples show that *zəfeʁehazerə-* with oblique-marked complements has the same meaning as with nominal arguments, namely 'prepare'. The verb *zəʁehazerə-*, which takes only clausal arguments, can also have the meaning 'intend, be going to', and it only takes complements with the adverbial case.

The potentialis with core cases is often used in the characteristic contexts for (de)verbal nouns in many languages, e.g. contexts of cultural events and activities rather than concrete situations:

(99) *sə-gʷ* *r-jə-hə-ʁ* [*təʁʷese-re* *qe.ɕʷe-**nə-r***]. (=31)
1SG.PR-heart LOC-3SG.A-carry-PST yesterday-ADJ dance-**POT-ABS**
'I liked yesterday's dances.'

Hence, the potentialis with core cases is used with the same meaning that complement-taking predicates have with nominal arguments. The use of this form is characteristic for (de)verbal nouns, and can be described as "cultural event or activity described by the nominalized verb". The semantic difference is sometimes very subtle, and cannot always be discerned even with the help of the broader context.

4.4 Opposition between the potentialis and the verbal stem with case markers

In sections 4.2–4.3 it has been demonstrated that the opposition between events and propositions in Adyghe is encoded by the case markers on the complement verb. Another issue to be explained is the distribution of the potentialis and the verbal stem in complementation. The choice between these forms is based on the temporal and modal meaning of the complement clause and the semantics of the complement-taking predicate.

In general, neither the verbal stem nor the potentialis show any restrictions on their temporal reference (with respect to the temporal reference of the matrix clause). Cf. examples (100) with the verbal stem and (101)–(104) with the potentialis. I distinguish four types of temporal reference:

(A) anteriority: the situation in the dependent clause precedes the situation in the matrix clause:

(100) a. [*s-šə-naḥə.ç'e* *čəja-**ʁe**-ç'e*] *me-g^{w}əʁe*. (=83a)
1SG.PR-brother-younger sleep-**PST**-INS DYN-hope
'She thinks he has fallen asleep {but he has not}.'

(B) simultaneity: the situation in the dependent clause is simultaneous with the situation in the matrix clause:

b. *se* *s-e-g^{w}əʁ^{w}e* [*we* *pçə* *wə-wəsə-ç'e*].
I 1SG.ABS-DYN-hope you lie 2SG.ABS-invent-INS
'I think that you might be telling me lies.'

(C) posteriority: the situation in the dependent clause follows the situation in the matrix clause:

c. [ə-*š'efə-**š'tə**-ç'e*] *w-e-g^{w}əʁ-a*?
3SG.A-**buy**-FUT-INS 2SG.ABS-DYN-hope-Q
'Do you think he's going to buy it? {I doubt it.}'

(D) gnomic interpretation of the situation in the dependent clause:

d. *se* *s-e-g^{w}əʁ^{w}e* [*we* *kaše-r* *ə-g^{w}* *r-jə-ḥ-ew*].
I 1SG.ABS-DYN-hope you porridge 3SG.PR-heart LOC-3SG.A-carry-ADV
'I hope that you like porridge.'

The verbal stem may include tense markers that encode the temporal reference of the dependent clause. Here (100a) and (100c) illustrate the markers of past

and future respectively, while present tense in (100b) and the gnomic reading in (100d) are signalled by the absence of any special marking[14].

Generally, the potentialis can encode all four types of temporal reference enumerated in (A)–(D), cf.:

(A) anteriority:

(101) [*vaze* *qʷətə-**ʁe-n**-č'e*] *s-e-š'əne*. (=20)
vase break-**PST-POT**-INS 1SG.ABS-DYN-fear
'I fear that the vase has broken' (during transportation).

(B) simultaneity:

(102) g*ʷəš'ə*ʔ*e* *pəte* *qə-w-e-s-e-tə* [*pçə* *sə-mə-wəsə-**n**-ew*].
word solid DIR-2SG.IO-OBL-1SG.A-DYN-give lie 1SG.A-NEG-invent-**POT**-ADV
'I give you my firm promise that I'm not lying.'

(C) posteriority:

(103) s*e* [*sə-če-**n**-ew*] *s-je-ž'ə-ʁ*. (=90)
I 1SG.ABS-run-**POT**-ADV 1SG.ABS-OBL-begin-PST
'I started running.'

(D) gnomic meaning:

(104) č'*ale-m* [*qe-ŝʷe-**nə-r***] *jə-č'as*. (=29b)
boy-OBL DIR-dance-**POT**-ABS POSS-favourite
'The boy likes dancing.'

In the case of anteriority (of the situation in the complement), the potentialis takes the past tense marker (101). The remaining three types of temporal reference are not encoded with special markers (102)–(104) and the exact interpretation is inferred from the context and from the semantics of the complement-taking predicate. There is a group of complement-taking predicates that require a future temporal interpretation of their complement, for example the verb 'begin' and verbs of potential situation ('try', 'intend' etc.) and causation and speech causation ('promise', 'permit' etc.). With these verbs the potentialis is always interpreted as referring to the situation that follows the situation in the matrix clause (103).

A number of complemenet-taking predicates exist that most often take complements with gnomic interpretation: for example, verbs with the semantics 'like'

14 In independent sentences, dynamicity markers occur in the present tense; they are absent in complement clauses with the verbal stem and the potentialis, cf. 3.1 and 3.2.

(104), 'love' and their antonyms, and evaluative predicates. However, these complement-taking predicates also allow for the posterior interpretation:

(105) [*pçə* *wə-wəsə-**nə**-r*] *deʁ^{w}-ep*.
lie 2SG.A-invent-**POT**-ABS good-NEG
a. 'It is not good to lie.'
b. 'If you lie now, it won't be good.'[15]

In (105) the complement clause with the potentialis can be interpreted both as posterior and gnomic meaning, and the exact interpretation is inferred from the context.

As has been demonstrated in Section 3.1, the potentialis can head not only subordinate, but also independent clauses, where it shows a peculiar distribution with the future marker. In independent clauses, the potentialis can only have future temporal reference; all other types of temporal reference arise in subordinate clauses only.

It should be noted that in Adyghe the temporal reference of the matrix clause does not impose strong restrictions on the choice of the tense form in the subordinate clause; most often it is not differentiated whether the tense of the subordinate clause encodes its temporal reference depending on the temporal reference of the matrix clause vs. on that of the speech act time (there is a pluperfect, however, it is not obligatory; there is no future-in-the-past).

Although all four types of temporal reference are attested with the potentialis, the last two types of temporal reference, posteriority and gnomic meaning, are by far the most frequent. In contexts of anteriority and simultaneity the potentialis occurs only marginally. Therefore, competition between the two forms mostly arises in contexts of posterior and gnomic interpretation. The use of the verbal stem in these two types of contexts is also restricted by the groups of complement-taking predicates. Hence, although both forms are acceptable in all four types of temporal reference, there is not a lot of variation in the choice between the potentialis and the verbal stem, because most complement-taking predicates impose restictions on this choice.

Below I describe the distribution of the verbal stem and the potentialis among those complement-taking predicates that can take one or both of these. Factive complement-taking predicates are not considered in this section. Most of the examples of the constructions discussed below have already been given in sections 4.1–4.2; here I only give examples that show deviation from the major pattern and have not been examplified in the previous sections.

15 In Adyghe, the 2nd person singular is used to introduce generic person (as in *One has to be patient*): this is why the generic-subject interpretation is possible in (105).

(A) Complement-taking verbs of immediate perception only take the verbal stem and the factive form; the potentialis is not attested with these verbs[16].

(B) Complement-taking predicates of evaluation can only take the potentialis and the factive form; the verbal stem is not attested.

(C) Complement-taking predicates of causation and potential situation, as well as the verb 'dream', require posterior temporal reference in their complement. These complement-taking predicates most often take the potentialis. The verbal stem with the future tense marker is marginal, although acceptable:

(106) [*sə-qə-b-de-g^{w}əš'ə?e-**š'tə**-ç'e*] *tha* *s-e-?^{w}e*.
1SG.ABS-DIR-2SG.IO-COM-talk-**FUT**-INS God 1SG.A-DYN-say
'I swear by the name of God that I'll talk to you.'

(D) Complement-taking emotive verbs behave differently with respect to the temporal reference of their complement. Verbs meaning 'love', 'like' and the verb 'bother' often require a gnomic interpretation in their complement. To encode this meaning, the potentialis is used. In other cases (e.g. *I liked how he danced yesterday*) only the factive form and the verbal noun have been attested. The construction using the verbal stem has not been attested with these complement-taking predicates.

The verbs 'want' and 'fear' and their synonyms most often take the potentialis where the complement has a posterior or gnomic interpretation. If the complement refers to a simultaneous or anterior situation, the verbal stem is preferable. However, the potentialis can also occur in these contexts (101).

(E) Complement-taking mental verbs and verbs of speech most often take the verbal stem if their complement refers to an anterior or simultaneous situation. However, the potentialis can also be used:

(107) [*kl'ejə-m* *oboj-xe-r* *ə-wəbətə-**ʁe-n**-ç'e*] *s-e-g^{w}əʁe*.
glue-OBL wallpaper-PL-ABS 3SG.A-catch-**PST-POT**-INS 1SG.ABS-DYN-hope/think
'I think that the glue has stuck to the wallpaper.'

If the complement refers to a posterior or gnomic situation, both the potentialis and the verbal stem are possible. The choice between the potentialis and the future in that case depends on the presence of modal meanings, which are characteristic for the potentialis (see Section 3.1).

16 It could be expected that the potentialis is acceptable in cases of the meaning of "cognitive perception" (*I see that you will not believe me*). However, such examples are absent in my sample.

The distribution of the potentialis and the verbal stem in the context of the different groups of complement-taking predicates is represented in Table 3[17]. I use round brackets for marginal constructions and the slash for where the two constructions are equally possible.

Table 3: Distribution of the potentialis and the verbal stem in complement clauses.

Groups of complement-taking predicates	Temporal reference of the complement clause (with respect to the temporal reference of the matrix clause)			
	anterior	**simultaneous**	**posterior**	**gnomic**
Perception verbs	–	verbal stem	–	–
Predicates of evaluation	–	–	potentialis	potentialis
Predicates of causation and potential situation and the verb 'dream'	–	–	potentialis (verbal stem)	–
Emotive verbs 'love', 'like', 'bother' etc.	–	–	–	potentialis
Emotive verbs 'fear', 'want' etc.	verbal stem (potentialis)	verbal stem (potentialis)	potentialis (verbal stem)	potentialis (verbal stem)
Mental verbs, verbs of speech	verbal stem (potentialis)	verbal stem (potentialis)	potentialis / verbal stem	potentialis / verbal stem

Thus, the distribution of the potentialis and the verbal stem can be explained on the basis of temporal and modal (and 'indirect speech act') meanings. The potentialis most often denotes a gnomic or posterior situation (with respect to the temporal reference of the matrix clause) that does not constitute the main assertion of the sentence. In other cases, the verbal stem is used.

5 Conclusions

The main complementation strategies in Adyghe are the factive form, the verbal stem with case markers, and the potentialis. There are also several further devices used in complementation: a paratactic construction, special forms used with phasal and modal predicates, a verbal noun of manner, a complementizer

17 I do not give the distribution of the factive form and the verbal noun in this table; only non-factive contexts are taken into account. For this reason the table does not include factive complement-taking predicates, such as 'be angry' and others.

derived from the verb of speech, and the relativization construction which is used to mark indirect questions.

Many complement-taking predicates in Adyghe allow non-subcategorized marking in complementation, which is attested with non-core cases on the verbal stem and the potentialis. These constructions meet the criteria for complement clauses and not adverbial clauses. I explain this peculiarity of the Adyghe complementation system in terms of the grammaticalization of the non-core cases as complementizers.

Thus, the number of complementation strategies in Adyghe is increased due to the grammaticalization of case markers in the function of complementizers. Non-core cases, the instrumental and the adverbial, are used in complementation to differentiate between the semantic types of complements. Events and irrealis propositions are marked with the instrumental case, while propositions are marked with the adverbial case. The polysemy of the instrumental case marker is unusual for complementizers in the languages of the world. However, it is only anomalous if we limit our research purely to the synchronic data. I suppose that this polysemy has a diachronic rationale. The two constructions have probably emerged from different sources and coincided in complement clauses. As discussed in Section 3.5.2, the "irrealis proposition" construction could originate in the antiresultative in *-ç'e*. The polysemy of irrealis and antiresultative is typologically widespread: as Plungian (2001: 13) shows, both meanings are often observed in markers of the pluperfect. I suggest that the "event" complement construction has developed from the instrumental case in *-ç'e*. It might have developed through the manner meaning of the instrumental (however, this meaning is rather marginal in Adyghe). The shift from manner towards event in the meaning of complementizers is attested in a number of languages, e.g. Slavic. Another diachronic scenario is the development of the event meaning from the verbal noun in *-ç'e* (which denotes manner and event with immediate perception verbs). However, this scenario is less probable since the two forms are differentiated morphophonologically. In sum, the exact path of grammaticalization of complements in *-ç'e* is unclear and requires further investigation.

Semantically, the encoding of clausal arguments in Adyghe is organized in the following way. The coreference pattern does not play a role in the choice of construction encoding the complement. The main relevant parameter is the semantic type of the embedded clause: factive, propositional, or eventive. However, unusual polysemy patterns occur: events are encoded by the same device as irrealis complements (instrumental case); manner complements are encoded in the same way as facts (factive form); the temporal reference of the complement clause is encoded by the tense markers on the verbal stem or potentialis. Topical

proposition complements are encoded with the factive form. Generic events are encoded by the same means as events proper.

The following table represents the distribution of the complementation strategies according to the semantics and temporal reference of the dependent clause (the distribution of the potentialis vs. future tense is presented in detail in Table 3).

Table 4: The choice of complementation strategy depending on the semantics and temporal reference of the complement clause.

Semantics of the complement clause	**The dependent situation is simultaneous with or precedes the situation in the matrix clause**	**The dependent situation follows the situation in the main clause or has gnomic interpretation**
Fact	factive form (with tense markers)	
Manner complements	factive form (with tense markers) / verbal noun in *-ç'e*	
Proposition	verbal stem (marginally: potentialis) with the adverbial case	potentialis / verbal stem with the adverbial case
Irreal or false proposition	verbal stem (marginally: potentialis) with the instrumental case	potentialis / verbal stem with the instrumental case
Event		
Generic event	n/a	potentialis with the instrumental case

It should be emphasized that the opposition between events and propositions is not encoded consistently; a number of eventive contexts are encoded by the factive form or by the verbal stem with the adverbial case (in the context of immediate perception verbs or phasal verbs). Thus, Adyghe makes a clearcut distinction between the context of fact and all other contexts; the event vs. proposition distinction is significant, but it is not always encoded consistently. It appears that the grammaticalization of non-core case markers in the function of complementizers is still in progress, or has stopped at this point.

Let us compare the Adyghe complementation system with that of Ossetic, which is considered in detail in (Serdobolskaya, this volume). The Ossetic system is sensitive to the opposition of coreferentiality patterns with complement-taking predicates of causation, speech causation and potential situation, emotive complement-taking predicates and others, thus demonstrating the same control pattern as is observed in Standard Average European languages. By contrast, Adyghe does not show any distinctions based on this parameter.

The same semantic parameters are relevant for the choice of complementation strategy in both languages. However, the exact "weight" laid on these parameters, and the polysemy patterns shown by the two languages, differ greatly. For Adyghe, the basic opposition is that of fact vs. non-fact (proposition/event). The event vs. proposition distinction is not encoded uniformly with all complement-taking predicates, as there are exceptions to the common pattern. Generic events are encoded by the same means as events proper. Irrealis complements take the same complementizer as eventives. Meanwhile, in Ossetic it is the event vs. non-event distinction that is most relevant; the fact vs. proposition distinction is not always encoded (which brings this system close to that found in Russian, where this differentiation is often reflected in the intonation pattern or deduced from context). There is a special device for marking irrealis complements and generic events. An unexpected polysemy pattern is observed with the conjunction *kʷəd*, which can encode event and potential / caused action in the future. A special device is used with the verb 'wait'.

However, some common features can be identified. In both languages facts are encoded by means of relativization: the relativization of the instrumental argument in Adyghe (i.e. the factive form), and the correlative construction in Ossetic. Belyaev and Serdobolskaya (in project) explain this feature of Ossetic in terms of the areal influence of the North-West Caucasian languages. For both Adyghe and Ossetic the information structure of the sentence is important for the choice of complementation strategy: the topicality of the complement triggers the choice of a special device reserved for the encoding of presupposed complements, which is the factive form in Adyghe and the correlative construction in Ossetic.

Appendix[18]. Distribution of complementation strategies

Complementation strategy	Factive form in *zere-*		Verbal stem		Potentialis in *-n*		
Matrix predicates	**acceptability**	**case marking/ postpositions**	**instrumental**	**adverbial**	**core cases**	**instrumental**	**adverbial**
faj- 'want'	–		–	–	–	+	+
ježe- 'wait'	–		–	–	?	+	+
Mental verbs							
ʂe- 'know'	+	ABS	–	+	–	+	+
š'əʁʷəpše- 'forget'	+	ABS	–	–	–	–	+
ʂʷeʂə- 'seem'	–		+/–	+	+/–	+/–	+
ɬəte- 'think, believe'	–		+	+	–	+	+
jenegʷəje- 'suppose'	?		+	+	–	+	+
š'eχʷə- 'think'	+	ADV	–	+	–	–	+
Predicates of emotion							
gʷəʁ(ʷ)e-[19] 'hope, think'	–		+	+	–	+	+
š'əgʷəʁ(ʷ)ə-[19] 'hope, think'	–		+	+	+	+	+
š'əne- 'fear'	–		+	–	–	+	+
š'əš'əne- 'fear'	–		+	–	+	+	+
jəçʼɑs 'like'	+	ABS	–	–	+	+	+
əgʷ rjehə- 'like'	+	ABS, INS	–	–	+	+	+
ʂʷə ɬeʁʷə- 'love'	+	ABS	–	–	+	+	+
χʷepse- 'dream'	+	OBL	–	–	+	+	+
jezeš'ə- 'annoy'	+	OBL, INS	–	–	+	+	+
gʷəʂʷe- 'rejoice'	+	OBL	–	–	–	–	–
fegʷəbžə- 'be angry'	+	INS, *paje* 'for'	–	–	–	–	–
Verbs of perception							
ɬeʁʷə- 'see'	+	ABS, ADV	–	+	–	–	–
zexexə- 'hear'	+	ABS	–	+	–	–	–

18 Notation in the Appendix: "+" means that a construction is acceptable, "–" that it is unacceptable; "+/–" that variation exists among native speakers; «?» marks insufficient information.
19 Both verbs can occur in two variants, *gʷəʁʷe-*, *š'əgʷəʁʷə-*, and *gʷəʁe-*, *š'əgʷəʁə-*, with no apparent semantic difference, cf. (Txarkaxo 1991).

Complementation strategy	Factive form in *zere-*		Verbal stem		Potentialis in *-n*		
Matrix predicates	**acceptability**	**case marking/ postpositions**	**instrumental**	**adverbial**	**core cases**	**instrumental**	**adverbial**
Speech verbs							
ʔʷe- 'say'	+	ABS	–	+	–	–	+
gʷəšʼəʔe (*je*)*tə-* 'promise', lit. 'give a word'	–		+/–	–	–	+/–	+
ʁegʷəʁe- 'promise, reassure'	+	INS	?	–	+	+	+
Predicates of potential situation or causation							
(f)jeʁeʂe- 'teach'	–		–	–	+	–	+
pələ- 'try'	+	OBL	–	–	+	–	+
zəʁehazərə- 'intend'	+	(POT)ADV	–	–	–	–	+
zəfeʁehazərə- 'intend'	+	(POT)ADV	–	–	+	–	+
hazər 'ready'	+	(POT)ADV	–	–	+	–	+
(*de*)*ʔepəʔe-* 'help'	–		–	–	+	–	+
fjədə- 'permit'	–		–	–	+	–	+
Evaluative predicates							
deʁʷə 'good'	+	ABS	–	–	+	+	+
(*fe*)*qjən* 'difficult'	–		–	–	+	+	+
psənçʼe 'easy'	–		–	–	+	+	+
tχaʁʷe 'pleasant'	?		–	–	+	+	–

Acknowledgment

This work has been supported by the RFH grants № 12-34-01345 and № 14-04-00580.

References

Abney, Steven P. 1987. *The English Noun Phrase in Its Sentential Aspect*. Cambridge: MIT dissertation.

Aissen, Judith L. 1974. Verb Raising. In *Linguistic Inquiry* 5(3). 325–366.

Apresyan, Juriy D. 1974. *Leksičeskaya semantika. Sinonimičeskiye sredstva jazyka* [Lexical semantics. Synonymic means of the language]. Moscow: Nauka.

Apresyan, Juriy D.. 1995. Problema faktivnosti: *znat'* i jego sinonimy [Problems of factivity: *znat'* and its synonyms]. In *Voprosy jazykoznaniya* 4. 43–64.

Arkadiev, Peter M., Juriy A. Lander, Alexander B. Letuchiy, Nina R. Sumbatova & Yakov G. Testelec. 2009. Vvedenie. Osnovnyje svedenija ob adygejskom jazyke [Introduction. Basic information on Adyghe]. In Testelec, Yakov G., Pjotr M. Arkadiev, Alexander B. Letuchiy & Nina R. Sumbatova (eds.). *Aspekty polisintetizma: očerki po grammatike adygejskogo jazyka* [Aspects of polysynthetic syntax: grammar sketches of Adyghe], 17–121. Moscow: RSUH.

Arutyunova, Nina D. 1988. *Tipy jazykovyx značenij: Ocenka. Sobytije. Fakt* [Types of linguistic meaning: Evaluation. Event. Fact]. Moscow: Nauka.

Asher, Nicholas. 1993. *Reference to abstract objects in discourse*. Dordrecht & Boston & London: Kluwer Academic Publishers.

Belyaev, Oleg I. & Natalia V. Serdobolskaya. In project. *Subordination in Ossetic and the typology of language contact: Syntactic replication without semantic influence*.

Boye, Kasper. 2012. *Epistemic meaning: A crosslinguistic and functional-cognitive study*. Berlin: Mouton de Gruyter.

Ershova, Ksenia A. 2011. Beslenejevskij masdar kak smešannaja kategorija [The Besleney masdar as a mixed category]. In Gerassimov Dmitriy V. (ed.) *Acta Linguistica Petropolitana VII (3)*, 289–295. Saint-Petersburg: Institute of Linguistic Studies.

Evans, Nicholas. 2007. Insubordination and its Uses. In Nikolaeva, Irina (ed.), *Finiteness*, 366–432. Oxford: Oxford University Press.

Gerassimov, Dmitriy. 2006. Complementation strategies in Circassian and Daghestanian. Paper presented at Morphosyntaxe des langues du Caucase, Paris, Inalco, 11–14 December.

Gerassimov, Dmitriy V. & Juriy A. Lander. 2008. Rel'ativizacija pod maskoj nominalizacii i faktivnyj argument v adygejskom jazyke [Relativization under the mask of nominalization]. In Plungian, Vladimir A. & Sergey G. Tatevosov (eds.). *Issledovanija po glagol'noj derivacii* [Studies in verbal derivation], 290–313. Moscow: Jazyki slavjanskix kultur.

Givón, Talmy. 1980. The binding hierarchy and the typology of complements. In *Studies in Language* 4 (4). 333–377.

Givón, Talmy. 1990. *Syntax: a Functional-Typological Introduction*, vol. 2. Amsterdam & Philadelphia: John Benjamins.

Kalinina, Elena Ju. 2002. *Nefinitnyje skazujemyje v nezavisimom predloženii* [Non-finite predicates as heads of independent sentences]. Moscow: Nasledije.

Kerasheva, Z. I. 1984. *Predloženija s finitnymi i infinitnymi glagolami v adygskix jazykax* [Clauses with finite and non-finite verbs in Circassian languages]. Tbilisi: Mecniereba.

Kerasheva, Z. I. 1995. *Izbrannyje trudy i statji* [Selected works and papers], vol. 1. Majkop: Adygejskoe respublikanskoe knizhnoe izdatelstvo.

Kimmelman, Vadim. 2010. Auxiliaries in Adyghe. http://www.uva.nl/over-de-uva/organisatie/medewerkers/content/k/i/v.kimmelman/v.kimmelman.html (2010)

Kiparsky, Paul & Carol Kiparsky. 1971. Fact. In Jakobovits, Leon & Danny Steinberg (eds.). *Semantics: An Interdisciplinary Reader*, 345–369. Cambridge: Cambridge University Press.

Knyazev, Michail Ju. 2009. Sentencialnyje dopolnenija v kalmyckom jazyke. In *Issledovaniya po grammatike kalmyckogo jazyka* (Acta Linguistica Petropolitana 4[2]), 525–581. Saint-Petersburg: Institute of Linguistic Studies.

Kumakhov, Mukhadin A. 1964. *Morfologija adygskix jazykov. Sinxronno-diaxronnaja xarakteristika* [Morphology of Circassian languages. Synchronic and diachronic characteristics]. Moscow, Nalchik: Kabardino-balkarskoje knizhnoje izdatelstvo.

Kumakhov, Mukhadin A. 1989. *Sravnitel'no-istoričeskaja grammatika adygskix (cherkesskix) jazykov* [Comparative grammar of Circassian languages]. Moscow: Nauka.

Kumakhov, Mukhadin A., Vamling C. 1998. *Dopolnitel'nyje konstrukcii v kabardinskom jazyke* [Complement constructions in Kabardian]. Moscow: Institute of Linguistics.

Kuznetsova, Julia L. 2009. Modal'nost v adygejskom jazyke [Modality in Adyghe]. In Testelec, Yakov G., Pjotr M. Arkadiev, Alexander B. Letuchiy & Nina R. Sumbatova (eds.). *Aspekty polisintetizma: očerki po grammatike adygejskogo jazyka* [Aspects of polysynthetic syntax: grammar sketches of Adyghe], 287–329. Moscow: RSUH.

Lehmann, Christian 1988. Typology of clause linkage. In Haiman, John & Sandra A. Thomson (eds.). *Clause Combining in Grammar and Discourse*, 181–225. Amsterdam, Philadelphia: John Benjamins.

Letuchiy, Alexander B. 2009. Affiksy benefaktiva i malefaktiva: sintaksičeskije osobennosti i krug upotreblenij [Affixes of benefactive and malefactive: syntactic properties and use]. In Testelec, Yakov G., Pjotr M. Arkadiev, Alexander B. Letuchiy & Nina R. Sumbatova (eds.). *Aspekty polisintetizma: očerki po grammatike adygejskogo jazyka* [Aspects of polysynthetic syntax: grammar sketches of Adyghe], 329–372. Moscow: RSUH.

Letuchiy, Alexander B. 2012. Ergativity in the Adyghe system of valency-changing derivations. In Authier, Gilles & Katharina Haude (eds.). *Ergativity, valency and voice* (Empirical approaches to language typology 48). Berlin, Boston, Mass.: Mouton De Gruyter.

Letuchiy, Alexander B. 2014. Russian clauses with *esli*: arguments or adjuncts? Factive or non-factive? Paper presented at Deutsche Gesellschaft für Sprachwissenschaft, workshop 'Clausal complementation and (non)-factivity', Marburg, 2014.

Malchukov, Andrej L. 2004. *Nominalization/verbalization: constraining a typology of transcategorial operations* (LINCOM Studies in Language Typology 8). München: LINCOM.

Miteva Polina S. 2009. *Konstrukcii s sentencial'nymi aktantami v ižemskom dialekte komi-zyr'anskogo jazyka* [Complement constructions in the Izhma dialect of Komi-Zyrian]. Moscow, ms.

Noonan, Michael. 1985. Complementation. In Shopen, Timothy (ed.). *Language Typology and Syntactic Description,* vol. 2*, Complex Constructions*, 42–140. Cambridge: Cambridge University Press.

Paris, Catherine & Niaz Batouka. 2005. *Dictionnaire abzakh (tscherkesse occidental)*, vol. 1: *Dictionnaire abzakh-français*. Paris: Louvain; Dudley: Peeters.

Plungian, Vladimir A. 2001. Antiresultativ: do i posle resultata. In *Issledovanija po teorii grammatiki* 1: *Grammaticheskiye kategorii* [Grammatical categories: Studies in the Theory of Grammar 1], 50–88. Moscow: Russkiye slovari.

Quer, Josep. 1999. Non-Logical *if*. In Satterfield, Teresa, Christina Tortora & Diana Cresti (eds.). *Current issues in Romance languages. Selected papers from the 29th linguistic symposium on Romance languages, Ann Arbor, 8–11 April 1999*, 241–255. Amsterdam & Philadelphia: John Benjamins.

Quer, Josep. 2008. Argument clauses and nominal features. *Lingue e linguaggio* 7(1). 93–110.

Rogava, Georgiy V. & Zejnab I. Kerasheva. 1966. *Grammatika adygejskogo jazyka* [Adyghe grammar]. Krasnodar: Krasnodarskoe knizhnoe izdatelstvo.

Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.

Ross, John R. 1967. *Constraints on variables in syntax*. Cambridge: MIT dissertation.

Serdobolskaya, Natalia V. 2009a. Vtoroje buduščeje, infinitiv, masdar i supin v adygejskom jazyke: argumenty za i protiv edinoj traktovki [Second future, infinitive, masdar and supine in Adyghe: pro and contra the unified interpretation]. In Testelec, Yakov G., Pjotr M. Arkadiev, Alexander B. Letuchiy & Nina R. Sumbatova (eds.). *Aspekty polisintetizma: očerki*

po grammatike adygejskogo jazyka [Aspects of polysynthetic syntax: grammar sketches of Adyghe], 454–498. Moscow: RSUH.

Serdobolskaya, Natalia V. 2009b. Semantika konstrukcij s sentencial'nymi aktantami v kunkinskom variante darginskogo jazyka [Semantics of complementation in Qunqi Dargwa]. http://www.kibrik.ru/content/pdf/Serdobolskaja.pdf (accessed 25 March, 2009)

Serdobolskaya, Natalia 2011. Grammaticalization patterns of the Adyghe instrumental case. In *Languages and Cultures in the Caucasus. Papers from the International Conference "Current Advances in Caucasian Studies", Macerata, January 21–23, 2010*, 515–539. München – Berlin: Verlag Otto Sagner.

Serdobolskaya, Natalia V. This volume. Semantics of complementation in Ossetic.

Serdobolskaya, Natalia V. & Alexander V. Motlokhov. 2009. Semantika konstrukcij s sentencial'nymi aktantami v adygejskom jazyke [Semantics of complementation in Adyghe]. In Testelec, Yakov G., Pjotr M. Arkadiev, Alexander B. Letuchiy & Nina R. Sumbatova (eds.). *Aspekty polisintetizma: očerki po grammatike adygejskogo jazyka* [Aspects of polysynthetic syntax: grammar sketches of Adyghe], 498–559. Moscow: RSUH.

Smeets, Rieks. 1984. *Studies in West Circassian phonology and morphology*. Leiden: The Hakuchi Press.

Sumbatova, Nina R. 2009. Kommunikativnaja struktura adygejskogo predloženija: perspektiva i fokus [Information structure of the sentence in Adyghe: perspective and focus]. In Testelec, Yakov G., Pjotr M. Arkadiev, Alexander B. Letuchiy & Nina R. Sumbatova (eds.). *Aspekty polisintetizma: očerki po grammatike adygejskogo jazyka* [Aspects of polysynthetic syntax: grammar sketches of Adyghe], 559–612. Moscow: RSUH.

Txarkaxo, Juriy. A. 1991. *Adygejsko-russkij slovar* [Adyghe-Russian dictionary]. Majkop: Adygejskoe knizhnoe izdatelstvo.

Yakovlev, Nikolay F. & Daud A. Ashxamaf. 1941. *Grammatika adygejskogo literaturnogo jazyka* [Grammar of Standard Adyghe]. Moscow, Leningrad: ANII kulturnogo stroitelstva.

Yanko, Tatiana E. 2001. *Kommunikativnyje strategii russkoj rechi* [Communicative strategies of the Russian speech]. Moscow: Jazyki slavjanskoj kultury.

Zalizniak, Anna A. 1990. O ponjatii "fakt" v lingvističeskoj semantike [On the concept of "fact" in semantics]. In Arutjunova, Nina D. (ed.) *Logičeskij analiz jazyka. Protivorečivost' i anomalnost' teksta* [Logical analysis of language. Contradictions and anomalies of discourse], 21–33. Moscow: Nauka.

Zalizniak, Anna A. 1992. *Issledovanija po semantike predikatov vnutrennego sostojanija* [Investigations into the semantics of inner state predicates]. München: Otto Sagner.

Zekokh, Učudžuk S. 1976. K probleme oslożnennogo predloženija v adygejskom jazyke [To the problem of non-finite subordination in Adyghe]. In *Strojenie predloženija v adygejskom jazyke* [The syntax of clause in Adyghe], 72–105. Majkop: Adygejskij NII ekonomiki, jazyka, literatury i istorii.

Zekox, Učudžuk S. 2002. *Adygejskaya grammatika* [Adyghe grammar]. Majkop: Adygeya.

Diana Forker

Complementizers in Hinuq

1 Introduction

This paper deals with complementation in the Nakh-Daghestanian language Hinuq. It additionally offers a short overview of complementation types in other languages of the same family. The focus is on the semantics, but wherever necessary a syntactic analysis will be provided. All Hinuq examples appearing in this paper have been collected by the author during fieldwork in Daghestan. Hinuq examples are not especially marked as such. For all examples from other Nakh-Daghestanian languages the sources are indicated together with the example sentences.

The structure of the paper is as follows: Section 2 provides an introduction to Hinuq grammar and its place within the Nakh-Daghestanian language family. In Section 3 I give a short introduction to all Hinuq complementizers and complementation types. In the subsequent Sections 4–6 the major complementation strategies are analyzed in detail. Section 7 is a short summary. Sections 8 and 9 deal with semantic choices in complementation and combinations of complementation types. Section 10 contains a synopsis of complement constructions built with other than purely verbal matrix predicates. In Section 11 the functions of complementizers outside of complement constructions are discussed. Section 12 presents an overview of complementation types and complementizers in Nakh-Daghestanian languages, and Section 13 contains the conclusion.

2 A short introduction to Hinuq and the Nakh-Daghestanian languages

The Nakh-Daghestanian languages represent the largest autochthonous language family of the Caucasus. The exact number of languages belonging to this family is unknown, but it can be estimated at around 40. The internal classification of the family has not yet been unanimously resolved. Thus, Figure 1 shows one of the possible classifications.

Nakh-Daghestanian (North-East Caucasian)

- Nakh branch
  - Chechen, Ingush, Tsova-Tush (Batsbi)
- Daghestanian branch
  - Avar-Andic subbranch
    - Avar
    - Andic
      - Andi, Botlikh, Godoberi, Karata, Akhvakh, Bagvalal, Tindi, Chamalal
  - Tsezic subbranch
    - Tsez, **Hinuq**, Khwarshi (incl. Inkhoqwari), Bezhta, Hunzib,
  - Dargi subbranch
    - Akusha/Standard Dargwa, Urakhi, Mugi, Tsudakhar, Gapshima-Butri, Mjurego-Gubden, Kadar, Muiri, Mehweb, Sirkhi, Amukh-Xuduc, Qunqi, Icari, Chirag, Kajtag, Kubachi, Ashti
  - Lak
  - Khinalug
  - Lezgic subbranch
    - Udi, Archi, Lezgian, Agul, Tabasaran, Tsakhur, Rutul, Kryz, Budukh

Figure 1: The Nakh-Daghestanian language family

Most of the Nakh-Daghestanian languages are spoken in southern Russia in the Caucasian mountains, but speech communities of e.g. Akhvakh, Khinalug, Udi, Lezgian or Kryz are also found in the neighboring regions of Georgia and Azerbaijan. The language at the center of this paper, Hinuq, is rather small. It is spoken by around 600 people living in the village of Hinuq in the Caucasian mountains near the Georgian boarder and in some other more recent settlements in the Daghestanian lowlands.

In order to facilitate the discussion of Hinuq complementation I will first present a short introduction to the basic facts of Hinuq syntax. Hinuq's morphology is basically agglutinative. It has a rather large inventory of finite and non-finite verb forms. The two largest groups of non-finite verb forms are participles mainly occurring in relative clauses and converbs predominantly used in adverbial clauses. Furthermore, the language has an infinitive and a verb form functionally equivalent to deverbal nouns and traditionally called 'masdar'.

Hinuq has a gender system with five genders that are used to mark agreement between nouns in the absolutive case and the majority of vowel-initial verbs. The gender/number prefixes marking agreement are given in Table 1. Usually the verb and the corresponding agreement trigger are part of the same clause, but in complement clauses it is possible that the matrix verb shows agreement with the absolutive argument of the complement clause. This is called 'long-distance agreement' (Section 3).

Table 1: Agreement prefixes

Gender/number	I	II	III	IV	V
SG	*Ø*	*y-*	*b-*	*y-*	*r-*
PL	*b-*	*b-/r-*	*r-*	*r-*	*r-*

The language is dependent-marking with a rich case inventory. The grammatical cases are absolutive, ergative, first and second genitive, dative and instrumental. Furthermore, there are 36 spatial cases. The verbal predicates can be divided into three basic types that correspond to the simple clause types of the language: (i) intransitive predicates having at least an S argument in the absolutive (including extended intransitive clauses with a further argument in a spatial case), (ii) transitive predicates with an ergative-marked A argument and an absolutive-marked O argument (including extended transitive predicates with a further argument in the dative or in a spatial case), and (iii) affective predicates with an experiencer in the dative and a stimulus argument in the absolutive. The most frequent word order in independent main clauses is SOV, but all other logically possible orders are also attested. In subordinate clauses the word order is more restricted and almost exclusively verb-final. For a recent grammar of Hinuq see Forker (2013).

Other Nakh-Daghestanian languages show properties very similar to the ones described for Hinuq, but some languages such as Lezgian or Agul do not have genders and lack agreement. Other languages as, for instance, Tsova-Tush, Dargi languages and Lak have person agreement in addition to gender/number agreement. The inventory of basic predicate types may also vary from language to language, but most if not all languages have a category of verbs with oblique subjects marked with a dative, an affective or a spatial case and including verbs such as 'know', 'want' and 'see'.

3 Complementation in Hinuq: An overview

3.1 Complementation markers and their uses

In Hinuq 'complementation' is encoded by a rather heterogeneous class of structures, which show little dedication to the given syntactic function. This is not the case in the marking of other complex sentences. Thus, adverbial clauses predominantly contain converbs and relative clauses mostly contain participles. Their

syntactic behavior is also quite mixed.[1] Hinuq has seven complement types that will be discussed in the following paragraphs:

1. zero marking (paratactic complement)
2. infinitive or purposive converb clause
3. complement clauses marked with the abstract suffix *-łi*
4. complement clauses marked with the quotative enclitic *=ƛen*
5. case-marked past participle or masdar
6. other non-finite verb forms (narrative converb, copula converb, habitual participle)
7. clause union

Hinuq does not have any particles similar to *that*, *if*, or *whether* used in English complement clauses. Young speakers very rarely use Russian complementizers; see examples (49) in Section 9. Interrogative pronouns are not considered to form a special type of complementizer, but their possible occurrences are discussed whenever important.

The focus of this paper will be on the types 1, 3, and 4, i.e. zero marking, the abstract suffix and the quotative enclitic because from a semantic point of view they represent the most interesting complementation types. In the following, all seven types will be shortly presented. Types 1, 3, and 4 will additionally be discussed in separate sections.

The zero-marking type is most commonly found with verbs of speech, but occurs also with a few other CTPs; cf. (1).[2]

(1) *xalq'i-la-y* *eƛi-yo* *zoq'e-n* *"b-iƛ'i-yo* *gom* *hadu"*
folk-OBL-ERG say-CVBI COP-UWPST III-go-CVBI COP.PRS.NEG this
'The people said, "It does not move."'

Complement clauses headed by the infinitive and complement clauses headed by the purposive converb occur with exactly the same CTPs, namely with (i) two propositional attitude verbs with the meaning 'think', (ii) commentative predicates, e.g. 'understand', 'agree', (iii) two verbs of knowledge or acquisition of knowledge, (iv) liking and fearing verbs, (v) manipulative predicates, (vi) modal predicates, (vii) achievement predicates, and (viii) two verbs with the meaning 'begin'. A number of these verbs such as modal verbs and many manipulative predicates do not occur with any other complement type apart from the infinitive/

1 See Thompson (2002) for a critique of complement constructions as a specific type of complex sentences.

2 Note that the absolutive is not indicated in the glosses. This means that all nominal expressions lacking a glossing for case are in the absolutive.

purposive converb type. Some examples of the use of the purposive converb and the infinitive with different complement taking predicates are presented in (2).

(2) a. *di-qo* *t'ek* [***cax-a*** / ***cax-ayaz***] *y-iq-iš-me*
1SG.OBL-AT book(IV) **write-INF** **write-PURP** IV-happen-PST-NEG
'I could not write the book.'

b. [*hayi-do* ***Ø-iƛ'-a***] *hałos=no* *pikru* *b-iq-iš=eƛ*
there-DIRC **I-go-INF** 3SG.M.GEN1=and thought(III) III-happen-PST=NARR
'He decided to go there.' (Lit. His thought happened to go there.)

c. *"de* [*dew-ho* ***nox-ayaz***] *razi* *gom=ƛen"* *eƛi-n*
1SG 2SG.OBL-ILOC **come-PURP** agree COP.PRS.NEG=QUOT say-UWPST
'She said, "I do not agree to marry you."'

d. *hałoy* *hago* *rek'we* [*haw* *bikore* ***b-uher-ayaz***] *kekir-no* *gom*
he.ERG that man(I) that snake(III) **III-kill-PURP** let-CVB COP.PRS.NEG
'He did not let the man kill the snake.'

The purposive converb is somewhat more commonly used than the infinitive. But this may eventually be due to the fact that almost all those verbs that have a stem-final vowel lack for morphological reasons the infinitive and use the purposive converb instead. The purposive converb suffix is *-(aya)z*, with *-z* being a cognate of the dative suffix *-z*.

Often the most prominent argument of complement clauses headed by the infinitive/purposive converb is controlled by either the most prominent argument of the matrix clause or another argument of the matrix clause. For instance, modal verbs require subject control (2a), and the manipulative predicate *kekir-* 'let, send' requires object control (2d). However, with other matrix verbs it is possible to dispense with control; cf. (3).

(3) *obu-z* *b-eti-n* [*uži-y* *mašina* ***b-ux-a*** / ***b-ux-ayaz***]
father-DAT III-want-UWPST son-ERG car(III) **III-buy-INF** **III-buy-PURP**
'Father wants his son to by a car.'

With a small number of predicates it is rather a possessor in the matrix clause that is coreferent with an argument of the complement clause (2b). The most prominent argument of the infinitival/purposive converb complement clause is generally omitted under referential identity with some main clause argument: (2a)–(2d).

The complement clauses formed with the infinitive and the purposive converb have irrealis meaning. Following the typology of Haspelmath (1989), they can be divided into:

- irrealis-prospective (e.g. with manipulative CTPs and desiderative CTPs, see example [2d])
- irrealis-potential (e.g. with modal CTPs, example [2a])
- realis-non-factive (e.g. with propositional attitude CTPs, example [2b])

Examples of complement clauses formed with the abstract suffix *-ɬi* and the quotative enclitic *=ƛen* are given in (4). The enclitic *=ƛen* is the only complementation marker in Hinuq that could be called a 'canonical complementizer' since its complements are finite. These complementation types are discussed at length in Sections 5 and 6.

(4) a. *hayɬoy* [*ƛerba-be* *b-aq'e-s-**ɬi***] *di-qo* *c'aɬer-iš-me*
he.ERG guest-PL HPL-come-RES-**ABST** 1SG.OBL-AT inform-PST-NEG
'He did not inform me that the guests came.'

b. [*haɬu-s* *maʔna* *se=**ƛen***] *uryezi* *Ø-iq-no* *hado* *uži*
this.OBL-GEN1 sense what=**QUOT** think I-happen-UWPST this boy(I)
'This boy thought, "What is the sense of this?"'

Only very few verbs take complements expressed with the past participle plus a case suffix, e.g. *boži -iq-* 'believe', *šak(ɬezi) -iq-* 'doubt' , *razi -iq-* 'be happy about', *ʔažaʔibɬi -uː-* 'wonder', and *ruq'eq' -iq-* 'be disappointed'. These CTPs consist of the Hinuq light verbs *-iq-* 'be, become, happen' or *-uː-* 'do' preceded by a loan word ultimately borrowed from Avar. The complements are nominalized clauses that are marked with case suffixes, which also occur on canonical NPs in the same position; cf. (5a) and (5b).

(5) a. *Maɦama* *boži* *Ø-iq-iš* *obu-ƛ'o*
Mahama(I) belief I-happen-PST father-SPR
'Mahama believed in his father.'

b. *Maɦama* *boži* *Ø-iq-iš* [*zones* *obu* ***Ø-aq'e-yoru-ƛ'o***]
Mahama(I) belief I-happen-PST REFL.SG.GEN1 father(I) **I-come-PST.PTCP-SPR**
'Mahama believed in the coming of his father.'

Some of these verbs can also take the masdar plus a case suffix with no difference in meaning between the two constructions:

(6) *de* *šakɬezi* *Ø-iqqo* [*zek* *ɣwede* *r-egi* ***r-iq-a-nu-ƛ'o***]
1SG doubt I-happen.PRS tomorrow day(V) V-good **V-happen-INF-MSD-SPR**
'I doubt that tomorrow will be a good day.' (i.e. good weather)

The masdar can also occur in certain complement of verbs of speech when talking about the topic of a conversation. In this construction, the clause expressing the topic is headed by a masdar plus the first genitive suffix.

(7) *de* *hayɬoqo* *xabar* *b-uher-iš* [*idu-do* ***Ø-aq'-a-nu-s***]
1SG.ERG 3SG.M.AT story(III) III-break-PST home-DIRC **I-come-INF-MSD-GEN1**
'I told him about (my) coming home.'

Furthermore, in the 'pretend' construction the matrix verb *-uː-* 'do, make' takes a complement containing a past participle marked with the first genitive suffix:

(8) a. [***toq-oru-s***] *r-uw-o!*
hear-PST.PTCP-GEN1 V-do-IMP
'Pretend to hear!'

b. *Madina-y* [*t'ek* ***t'ot'er-oru-s***] *r-u:-ho* *goł*
Madina-ERG book **read-PST.PTCP-GEN1** V-do-CVBI COP.PRS
'Madina pretends to read a book.'

Other non-finite verb forms used as complementation markers are the habitual participle, the narrative converb, and the converb of the copula. The latter occurs in embedded copula complements with CTPs that are verbs of cognition and knowledge; see the examples in (9). These complement clauses can be marked with the abstract suffix. If only the converb is used then the complements frequently imply some sort of uncertainty or they contain embedded questions (see Sections 5 and 8.1).

(9) a. [*haw* *šayt'an-i-š* *aqili* ***gołiš***] *hało-z* *bič'i* *r-iqqo*
she devil-OBL-GEN1 woman **COP.CVB** 3SG.M.OBL-DAT understanding V-happen.PRS
'He understands that she is a devil woman.'

b. [*dessu* ***gołiš*** *xan-i-š* *pikru*] *b-eq'i-r-an*
which **COP.CVB** khan-OBL-GEN1 thought(III) III-know-CAUS-INTFUT
'We will know what the khan's thought is.'

Only three phasal verbs make use of the narrative converb for their complement constructions: *łaq'e-* 'finish, end' (intransitive), *łaq'er-* 'finish, end' (transitive) and *-ičči-* 'be, stand, continue'; see (10) with the first one.

(10) *nesa:* [***Ø-iš-no***] *łaq'e-nos,* *hało* *ked-zo* *obu-y*
in.the.evening **I-eat-CVB** finish-ANT this.OBL girl-GEN2 father-ERG
hało-qo-r *toƛ-o* *saħ*
3SG.M.OBL-AT-LAT give-PRS 2,5kg
'In the evening until eating finishes, the girl's father gives him the measure container (of 2,5 kg).'

The habitual participle occurs in complements of immediate perception verbs (11a), of verbs of knowledge (11b) and of the phasal verbs *xece-* 'stop, let' and *-ičči-* 'stop'. The suffix of this participle, *-ƛ'o-s*, is morphologically complex. Both parts might be related to spatial case suffixes since Hinuq has a SPR-essive case *-ƛ'o* and an ablative case *-s*. But since there are no ancient documents of Hinuq or its relatives this cannot be proven. In (11a) the WH-word *deru* 'how' is obligatory. In (11b) and (11c) no WH-words are required. But the additional use of the abstract suffix to mark the complement clauses is possible and, in fact, more common than just the bare use of the habitual participle (see Sections 5 and 8.1). Example (11b) is remarkable because it shows that in Hinuq extraction of complement clauses is

possible: the pronoun *de* '1SG' either belongs to the first embedded clause or to the second embedded clause.[3]

(11) a. *ked-ez* *r-ik-o* [*deru* *meši-y* *bex* ***r-ac'-ƛ'os***]
girl-DAT V-see-PRS how calf-ERG grass(V) **V-eat-HAB**
'The girl sees how the calf is eating grass.'

b. [*q'idir* *Ø-iƛ'i-š*] *Ø-eq'i-š* *de* *hayło-z* [***Ø-uhe-ƛ'os***=*no*] *Ø-eq'i*
down I-fall-PST I-know-PST 1SG 3SG.M-DAT **I-die-HAB**=and I-know
'He knew that I fell down, so he will also know about my death.'

c. *me* *Ø-iči-yo-me* [***k'ošili:-ƛ'os*** *k'oɦlo-de*], *dew-qo-s* *de*
2SG I-stop-COND-NEG **play-HAB** ball(III)-ALOC 2SG.OBL-AT-ABL 1SG.ERG
b-iɣ-a *goł*
III-take-INF COP.PRS
'If you do not stop playing with the ball, I will take it away from you.'

Finally, clause union represents a distinct type of complementation and is only found with the CTP *-aq'e-* 'must'. Its complements can only be infinitives. The infinitival predicate does not head a clause of its own, but undergoes clause union. Thus, there is no embedded complement clause, but the construction is rather monoclausal.

(12) *Murad-i* *ze* *b-uher-a* *b-aq'e*
Murad-ERG bear(III) III-kill-INF III-must
'Murad must kill the bear.'

3.2 Semantic functions of complementation markers

With respect to the semantics of the complement, the complementation markers can be grouped together following Dixon's classification (2006):

- *potential type* ('refers to the potentiality of the subject of the complement clause becoming involved in an activity'):
  - infinitive/purposive converb
  - 'want' construction with the quotative enclitic *=ƛen*
  - clause union
- *activity type* ('refers to some ongoing activity, relating to its extension in time'):
  - zero marking
  - case-marked past participle or masdar

3 One reviewer suggested that maybe an alternative analysis is possible. Similar to the English sentence *I knew him to be a good footballer* with subject-to-object raising the pronoun *de* could eventually be part of the matrix clause. However, in contrast to the English example, the case marking of *de* does not help us to decide on this question. The pronoun *de* could be an S argument of 'fall' or 'die' as well as a stimulus argument of 'know'.

  - habitual participle
  - narrative converb
  - habitual participle plus abstract suffix *-ɬi* (with 'stop')
- *fact type* ('refers to the fact that something took place')
  - zero marking
  - abstract suffix *-ɬi*
  - quotative enclitic *=ƛen*
  - copula converb
- *speech act type* or *direct speech type* ('refers to a particular speech act')
  - zero marking with verbs of speech
  - verbs of speech plus the quotative enclitic *=ƛen*

The fourth type is actually not taken from Dixon who introduces only three types, but is based on the distinction of layers of clause structure by Functional Grammar (Hengeveld 1989: 130; Dik 1997: 93). It is relevant for CTPs that are verbs of speech (Sections 4 and 6).

Especially the second and the third type will be analyzed in more detail in this paper. Therefore, they warrant a few comments. The activity type has also been called "state of affairs" (Boye & van Lier 2009), "situational" (Haspelmath 1993) or "de re" reading (Frajzyngier & Jasperson 1991). It refers to the domain of reality, that is, the complement clause is a direct description of a situation, which is a spatial-temporal entity that can obtain or occur. It is therefore not epistemically evaluable and lacks a truth value.

Other labels for the fact type are "proposition" (cf. Boye & van Lier 2009) or "de dicto" (Frajzyngier & Jasperson 1991). This type belongs "to a semantic domain in which reference is made to the elements of speech" (Frajzyngier & Jasperson 1991: 135) because the complement clause in this case is represented as a fragment of speech (that may contain a description of a real event). This complement type is an abstract entity that is epistemically evaluable and hence bears a truth value.

In English, the distinction between activity type and fact type is manifested in the difference between *that*-complements and *ing*-complements:

(13) English
  a. *His eating broccoli was reported by the campus press.*
  b. *That he eats broccoli was reported by the campus press.*
  (Frajzyngier & Jasperson 1991: 142)

The complement clause in (13a) directly refers to the event itself whereas the complement in (13b) refers to a report about the event. In Hinuq, a similar difference can be obtained, for instance, by the use of case-marked participles/masdars

(activity type) or the quotative enclitic (fact type). See Section 8.4 for a detailed discussion of this distinction in Hinuq, and Section 12.3 for examples from other Nakh-Daghestanian languages.

The syntactic properties of complement clauses are not at the focus of this paper, but it is important to direct the attention of the reader to one remarkable property of Hinuq complement clauses. Many CTPs have agreement prefixes. What triggers the agreement is often an idiosyncratic property of the predicate (e.g. *-uǰ-* 'fear' is intransitive and its agreement is always triggered by its S argument which is the most prominent argument, i.e. the person who fears). However, a number of CTPs have two possibilities whose usages basically depend on the information structure of the sentence. These two possibilities are *local agreement* and *long-distance agreement*. Local agreement means that the CTP takes the *r-* prefix of gender v, a prefix which is also used in case of default agreement. In contrast, long-distance agreement means that the CTP agrees with the absolutive argument of the embedded clause which can be S, O or a stimulus. Long distance agreement is found in a number of other Nakh-Daghestanian languages including all Tsezic languages, Tsakhur (Kazenin & Testelec 1999), Godoberi (Haspelmath 1999), and Avar (Kibrik 2003: 459). Verbs that show long-distance agreement in these languages are 'know, want' and 'can'/'be able' (but not all these verbs in every language). The analyses of these constructions vary from language to language and sometimes even from verb to verb. Polinsky and Potsdam (2001) and Polinsky (2003) give a detailed analysis of long-distance agreement in Hinuq's closest relative Tsez. They argue that in Tsez, long-distance agreement is a topic-marking strategy for absolutive arguments, which, at least in transitive embedded clauses, are not typical topics. For Hinuq it rather seems that long-distance agreement is a pragmatic device for indicating focal arguments. It is used when the speaker wants to direct the attention of the hearer to the referent of the agreement trigger; see (14a). In contrast, sentences with local agreement are, so to say, neutral; none of the arguments in the embedded clause in (14b) is particularly salient.

(14) a. *Pat'imat-ez* *y-eq'i-yo* [*Madina-y* *t'ek* *y-ux-iš-łi*]$_{(v)}$
Patimat-DAT IV-know-PRS Madina-ERG book(IV) IV-buy-RES-ABST
'Patimat knows that Madina bought the BOOK.'

b. *Pat'imat-ez* *r-eq'i-yo* [*Madina-y* *t'ek* *y-ux-iš-łi*]$_{(v)}$
Patimat-DAT V-know-PRS Madina-ERG book(IV) IV-buy-RES-ABST
'Patimat knows that Madina bought the book.'

4 Zero marking

The zero (or paratactic) strategy means that two clauses containing finite verbs are juxtaposed without any overt marker expressing the relationship between them. They resemble asyndetically coordinated clauses, but one of the clauses contains a CTP. This means that the complement clause could also semantically and syntactically function as a main clause if it occurred on its own. This complementation type is possible with three types of CTPs: (i) verbs of speech, (ii) the immediate perception verbs 'hear' and 'see', and (iii) verbs of knowledge and acquisition of knowledge including the verb 'remember'. All verbs can also make use of other complement types, and zero marking is generally not the preferred strategy.

If the zero strategy is used with verbs of speech then the complement clause is regularly shifted to the right edge of the sentence. The complement clause is not marked with the quotative enclitic or phrase. Various verbs of speech make use of this strategy. Since the complement clause is headed by a finite verb, all TAM forms occurring in independent main clauses are allowed. For instance, (15a) illustrates the resultative present. Non-indicative moods like the imperative (15b) or interrogative clauses are also possible (15a).

(15) a. *hezzo* *axir-ƛ'o-do* *eser-no,* *"ɬu-y* *hibadu* *iši* *b-aq'er-iš*
then end-SPR-DIRC ask-UWPST who-ERG this apple(III) III-bring-RES
goł *hibału* *ked-ez?"*
COP.PRS this.OBL girl-DAT
'Then in the end (he) asks, "Who brought this apple to this girl?"'

b. *hayło* *xan-i* *qaƛ-o,* *"ey* *xan,* *mač'a* *maqo-r*
that.OBL khan-ERG call-PRS eh khan sword(V) outside-LAT
r-iɣ-o!"
V-bring.out-IMP
'The khan shouts, "Eh, khan, take your sword out!"'

These complements belong to the direct speech type. Nevertheless, this is a marginal way of expressing reported speech. The more frequent way of encoding reported speech is by means of the quotative enclitic, usually in combination with a quotative phrase (cf. Section 6). If the complement clause containing the quote is in its canonical preverbal direct object position, the zero strategy is not possible and marking with the quotative enclitic is obligatory.

Other verbs that use the zero strategy are verbs of immediate perception with activity complements. When occurring with the zero strategy, these verbs are used in their literal meaning. That means that the referent of the experiencer of *toq-* 'hear' must really hear some sound of the event referred to in the complement clause. For example, in (16b) the unexpressed experiencer hears the sound

of beating. It cannot mean 'hear' in the sense of 'get to know', or 'heard in a conversation'. The complement occurs in the canonical object position before the verb, (16a), or after the complement taking verb at the right end of the sentence, (16b).

(16) a. [*welosiped-ƛ'o* *toho-bito* *ked=no* *noxxo*] *y-ike-n...*
bicycle-SPR there-TRANS girl(II)=and come.PRS II-see-CVB
'When he also sees a girl on a bike coming from the other side...'

b. *toq-iš* [*zurmaqan-i* *zones=tow* *q'imu* *zokko*
hear-PST zurna.player-ERG REFL.SG.GEN1=EMPH head beat.CVBI
zoq'we-s=eƛ]
COP-PST=NARR
'(He) heard the zurna player was beating his own head.'

Note that as with the verbs of speech, the verbal head of the complement clauses can have exactly the same TAM forms as verbs of independent main clauses. For instance, in (16a) the verb in the complement clause takes the suffix of the simple present tense, and in (16b) it is marked with the simple past suffix followed by the narrative enclitic, thus marking the information conveyed by the complement clause as reported second hand information. Similarly, it is possible to have an interrogative form of the verb in the complement. This is a special verb form of the past that does not occur on non-finite predicates. Moreover, this construction only marginally allows for local agreement. All corpus examples with the zero strategy exhibit long-distance agreement as in (16a), but in elicitation local agreement is also available.

The last group of CTPs that allows for zero-marked complement clauses is verbs of knowledge and acquisition of knowledge including 'remember'. These verbs behave slightly different from the other two groups. First of all, it seems that zero-marked complements of these matrix predicates occur more frequently in clause initial position as a kind of topic, see (17a), than sentence-finally.

(17) a. [*nišƛa* *mix* *zoq'e-n*] *b-eq'i-yo* *gom*
which time(III) COP-UWPST III-know-CVBI COP.PRS.NEG
'(I) don't know what time it was.'

b. *diž* *c'aɬ-iš* [*xan* *Ø-aq'e-s*]
1SG.DAT get.to.know-PST khan(I) I-come-PST
'I got to know that the khan came.'

It can be hard to distinguish between zero-marked complement clauses and parentheticals, i.e. words, phrases, or sentences which interrupt a sentence and which bear no syntactic relation to that sentence at the point of interruption, e.g. *You will, I think, have to buy a new car*. The Hinuq corpus on which this study is based almost exclusively consists of monologues (fairy tales, legends, auto-

biographical narrations, procedural texts, etc.). Therefore, it does not have many examples of parenthetical usages of CTPs that seem to be extremely common in everyday conversation (Thompson 2002). Almost all examples occurring in my corpus are with the verb *rok'ƛ'o goł* 'remember'; see (18).

(18) *hoboy* *hezzo* *iškola-do* *kekir-oƛ'o=n,* *rok'ƛ'o* *goł* *diž,*
then later school.in-DIRC send-SML=and heart.SPR COP.PRS 1SG.DAT
diž *iškola-do* *Ø-iƛ'-a* *Ø-eti-yo* *zoq'we-s-me*
1SG.DAT school.in-DIRC I-go-INF I-want-CVBI COP-PST-NEG
'Then later when (my mother) sent me to school, I remember, I did not want to go to school.'

In elicitation speakers claim that complements of 'remember' must be marked with *-łi* (Section 5) or *=ƛen* (Section 6) if the intended meaning is 'remember that'. Or they must be marked with the infinitive/purposive converb if the intended meaning is 'remember to'.[4] However, in autobiographical narrations I found a number of examples where what seems to be the complement clause is unmarked, and the apparent matrix clause 'I remember' occurs in its own intonational phrase either breaking up what could be considered the complement or following it. Both word orders are untypical for most complement clauses, including zero-marked ones (see Section 4). In examples like (18) the phrase 'I remember' bears no clear syntactic relationship to the preceding and/or the following clause. In terms of Thompson (2002), the verb 'remember' in such examples does not constitute the main point of the utterance. Since the sentence is part of an autobiographical narration, it is clear that everything that the speaker recounts is part of his memories. From time to time he interrupts his narration with the fixed formulaic phrase 'I remember' which serves as an epistemic frame for his narration. It seems that in such examples the phrase 'I remember' has undergone grammaticalization into a 'grammatical and secondary CTP' (Boye & Harder 2007), although it is clearly less grammaticalized than the matrix verb 'say' in reported speech (see Section 6).

From a formal point of view, zero-marked complements show the least integration with the matrix predicate leading to the extreme case of parenthetically used CTPs with no integration. The complements denote propositions (fact type), state of affairs (activity type), or particular speech acts (direct speech type). The zero-marked complements denoting propositions or direct speech have independent time reference and allow for all verb forms that are available in main clauses, including non-firsthand evidential markers such as the unwitnessed past in (17a) or imperatives as in (15b) (see Sections 8.1 and 8.2 for more examples).

4 Thus, when I checked an example similar to (18) with the speaker, he said that it is better to add the abstract enclitic to the complement clause.

5 The abstract suffix *-ɬi*

This suffix is a loan from the related language Avar. Its use in Avar is similar to its use in Hinuq: *-ɬi* is used (i) to derive abstract nouns from nouns and adjectives, and (ii) to optionally mark complement clauses, for instance, complements of the verb 'know', as in (19).

Avar

(19) *dida* *ɬana* [*soncicin* *hey* *zavodalda* *yik'inč'o-**ɬi***]
1SG.LOC knew yesterday she fabric.LOC not.was-**NMLZ**
'I knew that she was not in the fabric yesterday.'
(Alekseev & Ataev 1998: 110–111)

In Hinuq, the abstract suffix *-ɬi* has six different functions:

1. Formation of abstract nouns from nouns and very few adverbs, e.g. *xan-ɬi* 'kingdom' < *xan* 'king', *bac'ad-ɬi* 'cleanness' < *bac'ad* 'cleanly'. Hinuq also contains a number of nouns ending with *-ɬi* which have been borrowed as a whole from Avar because the corresponding base words do not exist in Hinuq, e.g. *bercin-ɬi* 'beauty', but there is no word *bercin* in Hinuq.
2. Marking of complement clauses. This function will be looked at in more detail below.
3. Marking of the topic of a conversation or narration, e.g.

   (20) *Ħapiz* *ʔumar* *essu-**ɬi*** *ese-s*
   Hapiz Umar brother-**ABST** talk-PST
   '(I) talked about brother Hapiz Umar.'

4. Marking X in phrases like 'X turns into Y' (21). In this function it may optionally be followed by the suffix *-žo*.

   (21) *hału* *ked-i* *zon-**ɬi*** *b-uː-ho* *arxi,*
   this.OBL girl-ERG REFL.SG.OBL-**ABST** III-do-PRS ditch(III)
   gulu-za-ɬi *r-uː-ho* *ɬe*
   horse-OBL.PL-ABST V-do-PRS water(V)
   'The girl turns herself into a ditch and the horses into the water.'

5. Formation of a simultaneous converb (by adding *-ɬi* to the infinitive or with some verbs directly to the stem):

   (22) [*Ø-ox-a-**ɬi***] *haze* *łono=n* *uži-qo=gon* *b-aši-yo* *hayłos*
   I-leave-INF-**ABST** those.OBL three=and boy-AT=TOP III-find-CVBI 3SG.M.GEN1
   ƛ'oq'on
   hat(III)
   'While he is leaving, the three boys find his head.'

6. Marking of complements of the postposition *ʔolo* 'because of, for the sake of'. In this function the abstract suffix may be followed by another suffix *-žo* or replaced with the suffix *-ƛi*:

(23) *xexza-**ƛi*** *ʔolo* *eli* *aƛ-a-do* *nox-iš*
child.OBL.PL-**ABST** because.of 1PL village-in-DIRC come-PST
'Because of the children we came to the village.'

The first three functions and probably also the fourth and eventually the fifth function are related because they all involve abstract nominals, either as genuine parts of the nominal lexicon, or in a more theoretical sense: complements can be analyzed as nominalized propositions that occur in a position where nouns usually occur (see below). The topic of a conversation and the item into which another object or animate being has been turned can be viewed as abstract nominalized objects. Finally, it might be that for the fifth function a subordinate clause is transformed into a nominal that serves as the background for the main clause action, but the verbal head of the subordinate clause lacks the formal properties of nouns (e.g. case marking). However, in this paper I will only analyze the employment of *-ƛi* as a marker of complement clauses.

When used in complement clauses *-ƛi* must occur on the verb that constitutes the head of the complement clause, and the position of this verb is obligatorily clause-final. The majority of CTPs that make use of the suffix *-ƛi* can be characterized as factive verbs and further divided into several subgroups. The largest group are verbs of knowledge and acquisition of knowledge such as *-eq'i-* 'know', *c'aƛ-* 'get to know', *bič'i -iq-* 'understand', or *c'aƛer-* 'let know'. The first three verbs belong to the valency class of affective verbs taking an experiencer in the dative and a stimulus argument in the absolutive case (see Section 2). Other CTPs employing *-ƛi* are two achievement predicates (*šuƛe-* 'forget', *rok'ƛo(r) -aq'e-* 'remember'), immediate perception predicates (*-ike-* 'see', *toq-* 'hear'), evaluative predicates such as 'it is good that', and the verbs *-ace-* 'hate' and *xece-* 'let, stop'. All verbs that occur with this complementation marker also have other possibilities of forming complement clauses such as the zero strategy, the infinitive/purposive converb, the quotative enclitic or the habitual participle.

The verb forms in complement clauses marked with *-ƛi* can be of three formal types: (i) converb of the copula (for embedded copula clauses), (ii) resultative participle (for complement clauses with past time reference), or (iii) habitual participle (for complement clauses with present or future time reference or for complement clauses lacking explicit time reference). Finite verb forms are not allowed. Relevant examples are presented in (24a)–(24f), but also in (41a) and (41b) below with immediate perception predicates in the matrix clause.

(24) a. *diž* [*me* *Ø-iq-iš* ***gołiš-łi***] *Ø-eq'i-n* *zeq'we-s-me*
1SG.DAT 2SG I-become-RES **COP.CVB-ABST** I-know-CVB COP-PST-NEG
'I did not know that you (masc.) were born.'

b. *diž* *neteqen* *šuƛ'e-me* [*eli* *cadaq* *kino-mo-ł-er* *b-exna:-ho*
1SG.DAT never forget-NEG 1PL together film-OBL-CONT-LAT HPL-go-CVBI
zoq'we-s-łi]
COP-RES-ABST
'I will never forget that we went together to the movies.'

c. *hayłu-z* *rok'ƛ'o* *r-aq'-o* [*essu-y* *t'ek* ***t'ot'er-iš-łi***]
3SG.F.OBL-DAT hear.SPR V-come-PRS brother-ERG book **read-RES-ABST**
'She remembers that the brother read the book.'

d. *diž* *r-ac-o* [*hayłoy* *ʔaraq'a* ***ga:-ƛ'os-łi***]
1SG.DAT V-hate-PRS 3SG.M.ERG vodka **drink-HAB-ABST**
'I hate that he drinks vodka.'

e. [*ked-i* *t'ek* ***t'ot'er-iš-łi***] *r-eg* *goł*
girl-ERG book **read-RES-ABST** V-well COP.PRS
'It is good that the girl read the book.'

f. *me* [*t'ek-mo-za-ł* ***toqli:-ƛ'os-łi***] *xec-o!*
2SG.ERG book-OBL-OBL.PL-CONT **rummage-HAB-ABST** stop-IMP
'Stop rummaging with the books.'

As can be seen from the examples in (24), complement clauses marked with *-łi* normally describe situations that exist independently of the actions or the states described in the matrix clause. They are part of what is asserted and have realis modality. The information conveyed in the complement clause can be backgrounded material (24a) or it can be new in the discourse context (25). Their time reference is independent of the time reference of the CTP. This can be illustrated by (25): this example contains two complement clauses embedded within the same CTP. The first embedded clause has past-time reference due to the resultative participle on the clausal head; the second complement clause has future-time reference because the verbal head has the habitual participle suffix attached to it.

(25) *hezzo* *hału-z* *bič'i* *r-iqqo* [*zo* *q'orol-a-y*
then 3SG.F.OBL-DAT understanding V-happen.PRS REFL.SG widow-OBL-ERG
y-aƛ'ir-iš-łi] [*zonde-r* *xoddo* *t'ok'aw* ***Ø-aq'e-me-ƛ'os-łi***]
II-betray-RES-ABST REFL.SG.ALOC-LAT husband(I) anymore **I-come-NEG-HAB-ABST**
'Then she understands that the widow had betrayed her, that her husband would not come back to her.'

To sum up, these complement clauses can be said to be less integrated and less dependent than, for instance, complements headed by the infinitive/purposive converb, but more integrated and more dependent than zero-marked complements. They preserve their argument structure. However, formally, they are clearly dependent clauses since they contain non-finite verb forms and the suffix *-łi*. Neither of these formal devices is found in independent main clauses. With

regard to their semantics, it is clear that they denote propositions that usually have a realis meaning. The propositions are often epistemically neutral, but can occasionally express uncertainty (see Section 8.1 for a more detailed discussion).

As has been mentioned above, depending on the CTP, complement constructions of the *-łi* type can show local agreement, see (24c)–(24e) and (25), long-distance agreement (24a) or no agreement at all, see (24b) and (24f). The examples with local agreement can be taken as argument for claiming that complement clauses marked with *-łi* are in fact nominalized clauses. As shown above, *-łi* is used as a derivational suffix for forming abstract nouns most of which belong to gender V (prefix *r-*). Similarly, complement clauses with *-łi* trigger the prefix *r-* on the CTP in case of local agreement. However, also infinitival/purposive converb complements can trigger local agreement marked with the prefix *r-* on the matrix predicate.

6 The quotative enclitic *=ƛen*

The major function of the quotative enclitic is to mark reported speech. In addition, it also appears with other types of CTPs, namely propositional attitude verbs, some verbs of knowledge and acquisition of knowledge, some liking and fearing verbs, and a few other verbs. Furthermore, the enclitic can be used in a complementation-like construction lacking CTPs (see below). Finally, the enclitic can express hearsay evidentiality.

For other Tsezic languages that have a cognate quotative enclitic (Khwarshi *ƛin/ƛun*, Tsez *ƛin*, Hunzib *ƛe*, and Bezhta *ƛo*) it has been proposed to analyze the enclitic as a grammaticalized form of the narrative converb form of the verb 'say' (see e.g. Khalilova 2009: 472 on Khwarshi). In fact, in many languages quotative particles are diachronically derivable from verbs like 'say' (Aikhenvald 2011). But although the Hinuq narrative converb of 'say', *eƛin*, resembles the enclitic *=ƛen*, the stem final vowel of the verb is *-i* and not *-e*.

The main function of the quotative enclitic is the marking of reported speech. It usually appears on the last word of the clause representing the speech act, see (26a), but occasionally it can occur more than once, on the last word and on some other word (often the first word) belonging to the same clause, see (26b). It can occur on all parts of speech: nouns, pronouns, proper names, WH-words, adjectives, adverbs, etc., but the preferred hosts are verbs which usually appear in clause-final position.

(26) a. *“sira* *me* *hayru* *wed* *r-uː-ho”=**ƛen*** *waʔza* *b-uː-s*
why 2SG.ERG so matter(V) V-do-PRS=**QUOT** reminding(III) III-do-PST
hayɬo-qo-r
3SG.M.OBL-AT-LAT
‘“Why do you do such things?” (the mullah) asked him.’

b. *hału* *očordiyu* *aqila-y=gon* *“me”=**ƛen*** *eƛi-yo,*
this.OBL old woman.OBL-ERG=TOP 2SG=**QUOT** say-PRS
“aƛaza-r=no *Ø-exnaː-n”=ƛen* *eƛi-yo*
village.PL.in-LAT=and I-go-UWPST=QUOT say-PRS
‘“You”, the old woman says, “went to the villages.”’

Not only verbs of speech and all manipulative predicates referring to speech acts (e.g. ‘command’, ‘request’), but even verbs that usually do not take complements can have a complement clause marked with the quotative enclitic; cf. (27).

(27) *obu-y* *eluqo-r* *wasi* *b-uː-n* *zeq’we-s* *“ɬora=n*
father-ERG 1PL.AT-LAT testament(III) III-do-CVB COP-PST three.OBL=and
ɣwede-ł *hayɬo-zo* *sud-a-ho* *axranƚi* *r-uw-o”=**ƛen***
day-CONT 3SG.M.OBL-GEN2 grave-OBL-ILOC watch(V) V-do-IMP=**QUOT**
‘Our father said (lit. made) in his testament to us to guard his grave for three days.’

The use of *=ƛen* alone is enough to indicate such complements. But often the enclitic appears in combination with the narrative converb of the verb *eƛi-* ‘say’. This converb then is used as a further particle indicating the reported speech construction, and not as a matrix verb taking itself a complement. Yet it is not always possible to precisely identify which occurrence of *eƛin* represents the grammaticalized particle and which the full lexical verb taking a quote as its complement. Thus, it is possible to analyze *eƛin* in such contexts as a kind of parenthetical CTP (cf. Section 4) that proceeds on its grammaticalization path from a full CTP to a mere discourse particle marking reported speech. Its position varies and it can occur more than once marking nevertheless only one utterance, just like the quotative enclitic; see (28a). It cannot be modified by a manner adverbial, (28b), which indicates its parenthetical status.

(28) a. *boc’e-y* ***eƛi-n*** *hayɬu-qo-r* *“me* *de* *b-uher-an”=**ƛen***
wolf-ERG **say-UWPST** 3SG.OBL-AT-LAT 2.SG 1SG.ERG III-kill-INTFUT=**QOUT**
eƛi-n, *“de* *nekwe-s* *goɬ”=ƛen* ***eƛi-n,*** *“diž* *b-ac’-a*
say-CVB 1SG starve-RES COP.PRS=QUOT **say-CVB** 1SG.DAT III-eat-INF
*b-eti-n”=**ƛen*** ***eƛi-n***
III-want-UWPST=**QUOT** **say-CVB**
‘The wolf said to it (the fox): “I will kill you, I am hungry, I want to eat.”’

b. *de* *hayɬoqo* *eser-iš* *“ni-š* *me* *Ø-aq’e-y”=**ƛen*** (*ɣeɣe
1SG.ERG 3SG.M.AT ask-PST where-ABL 2SG I-come-Q=**QUOT** quietly
Ø-iči-n) ***eƛi-n***
I-be-CVB **say-CVB**
‘I asked him (*quietly) where he came from.’

The complement clause, especially when it is quite short, can occur in the canonical object position immediately before the verb of speech, but shifting it to the left before the matrix clause or to the right after the matrix clause is also frequently attested.

In the most extreme case, a quote can just be indicated by the use of *=ƛen* without any accompanying particle *eƛin* or a matrix verb of speech. This means that CTP ellipsis is possible; see (29). However, this type of ellipsis is restricted to reported speech; if the CTP is not a verb of speech it cannot be omitted.

(29) *“haɬu-s* *hunar=e* *debez* *b-iker-a* *b-eti-n”=**ƛen***
3SG.F.OBL-GEN1 ability(III)=Q 2SG.DAT III-show-INF III-want-UWPST=**QUOT**
b-ux-no *Malla* *Rasadan-i* *ɣwadi=n*
III-take-CVB Mullah Nasredin-ERG crow=and
‘“You want to see her ability?” asked Mullah Nasredin, taking the crow.’

Reported speech constructions in Hinuq do not use a special verb form in the quote. There is no tense shift and more generally no shift in the spatial or temporal deixis. Reported commands, wishes, and questions are marked just like in normal main clauses. These complement clauses are of the fourth semantic type, the direct speech type. All verb forms of the non-indicative mood or with a non-indicative modality may occur in quotes of reported speech events just like in normal main clauses; such verb forms include imperatives, e.g. (30), prohibitives or optatives, and interrogative verb forms. The only difference that may distinguish quotes from ordinary main clauses is the occurrence of reflexive pronouns. Both speaker and addressee are most frequently encoded with first and second personal pronouns in the quote; see (28a) and (28b). However, it is possible to use reflexive pronouns in the quote that can only refer to the speaker, never to the addressee; see (30).

(30) *exo-qo-r* *Malla* *Rasadan-i* *eƛi-yo* *“zo*
herdsman-AT-LAT Mullah Nasredin-ERG say-PRS REFL.SG
q’iloza-qo-s *Ø-ekir-o”=ƛen*
board.OBL.PL-AT-ABL I-unbind-IMP=QUOT
‘Mullah Nasredin says to the herdsman: “Unbind me!”’

Apart from reported speech, *=ƛen* is used as a complementizer with the following CTPs:

1. Propositional attitude verbs such as ‘seem’, ‘think’, ‘believe’, ‘doubt’ and ‘be sure’; cf. (31a). The verb *-ese-* ‘be probable’ occasionally appears in these complements, expressing doubts on part of the subject referent (i.e. the S, A or EXP argument of the matrix verb) concerning the truth of the proposition in the complement clause.

2. Verbs of knowledge and acquisition of knowledge (with the exception of *-eq'i-* 'know'); cf. (31b). Furthermore, the verb *toq-* 'hear' can employ this strategy when it is used in the sense of 'get to know (from a conversation)'; cf. (31c).
3. Two achievement predicates, namely 'try, imagine' and 'remember'; cf. (31d).
4. Some liking and fearing verbs such as 'hope', 'fear', 'be angry' and 'hate' (31e).

(31) a. [*r-ič'-iš* *gadi* *r-iži-yo=**ƛen***] *qeba:-s=eƛ* *nartaw-za-z*
v-fill-RES barrel(v) v-take-PRS=**QUOT** seem-PST=NARR giant-OBL.PL-DAT
'The giants thought that he would bring a filled barrel.'

b. [*hago* *čanaqan* *Ø-aq'e-n=**ƛen***] *c'aɬ-oƛ'o* *hayło* *xan-i*
that hunter(I) I-come-UWPST=**QUOT** get.to.know-SML that.OBL khan-ERG
kekir-ho *wazir* *čanaqan* *eƛ-ayaz*
send-PRS vizier hunter say-PURP
'When the khan gets to know that the hunter came, he sends the vizier to call the hunter.'

c. *seda* *zaman-a-ɬ* *toq-no* [*seda* *uži-ž* *pulanaw* *xan-es*
one.OBL time-OBL-CONT hear-UWPST one.OBL boy-DAT certain khan-GEN1
ked *goɬ=**ƛen***] *eƛi-n*
girl COP.PRS=**QUOT** say-CVB
'One time one boy heard that a certain khan has a daughter.'

d. [*hibay* *bečedaw* *dawla=n* *b-ux-no* *halmaɣ* *Ø-aq'e=**ƛen***]
such rich wealth(III)=and III-keep-CVB friend(I) I-come=**QUOT**
zurmaqan-zo *neteqen* *rok'ƛ'o* *r-aq'e-s-me=eƛ*
zurna.player-GEN2 never heart.SPR V-come-PST-NEG=NARR
'The zurna player could never imagine that his friend would come bringing such wealth.'

e. *di* *kul* *goɬ* [*zek* *ɣwede* *r-egi* *r-iq=**ƛen***]
1SG.GEN1 hope v tomorrow day(v) v-good v-become=**QUOT**
'I hope that tomorrow will be a good day.'

As with reported speech constructions, verbs in these complement clauses are finite and have thus independent temporal reference. In general, they express all those grammatical categories that are expressed in true independent clauses. They can be realis or irrealis, depending on the CTP and on the verb form employed in the complement clause. Irrealis complements are especially frequent when the embedded complement is expressed as a TAM form that can have future time reference, e.g. the simple present or the general tense, see (31a) and (31d), or if the complement clause contains the epistemic-modal auxiliary *-ese-* 'be probable, be possible', see (32). The S, A, EXP, or addressee argument of the CTP can but does not have to be identical to the S, A, or EXP argument in the subordinate clause.

Complements marked with the quotative enclitic can contain modality markers, which contradicts Frajzyngier's (1995) claim that complementizers, being primarily modal, cannot co-occur with other modality markers in the

complement clause (see van Lier & Boye 2010 for more counterexamples to this claim).

(32) [*me* *rek'we* ***Ø-ese=ƛen***] *qeba:-ho* *zoq'e-s*
2SG person(I) **I-be.probable=QUOT** think-CVBI COP-PST
'I thought that you were probably a person.'

Finally, there is a last type of complementation-like construction that bears the quotative enclitic. The meaning of this construction is 'want to' or occasionally 'in order to', i.e. irrealis. Usually, in Hinuq such a meaning is rendered by means of a matrix verb such as *-eti-* 'want' and a complement clause headed by the infinitive or the purposive converb (cf. Section 3.1). However, it seems that a new construction with similar meaning is currently developing. In this construction, the situation that is described as wanted or as the goal of an action is expressed by a verb in the intentional future to which *=ƛen* is added and there is no matrix verb; see the examples in (33).[5]

(33) a. *hałoy* *zoq'e-n* ***b-ux-an=ƛen*** *hes* *geni*
3SG.M.ERG COP-UWPST **III-take-INTFUT=QUOT** one pear(III)
'He wanted to take one pear.'
b. *hayłoł* *hezzo* [*hago* ***Ø-ašir-an=ƛen***] *hadbe* *q'ono=no* *k'oƛe-n*
3SG.M.CONT after 3SG.M **I-catch-INTFUT=QUOT** these two=and jump-UWPST
'These two run after him in order to catch him.'

The intentional future is normally only used with first persons in assertions, never with second or third persons (see Forker 2013: 201–204 for examples in simple clauses). It expresses the intentions of the speaker; what s/he wants to do. Yet in the complementation-like construction exemplified by (33a) and (33b) there is no person restriction anymore. Thus, we have to explain why the intentional future can be used in these examples. Furthermore, we need to account for the desiderative or purposive semantics that (33a) and (33b) convey considering that they do not contain appropriate CPTs. My suggestion is that this construction has developed from complement clauses such (31e), or (34).

(34) *hałoˑ-s* *xiyal* *b-iq-o* [*b-ašir-no* *besuro,* *xan-i-ž*
3SG.M.OBL-GEN1 wish(III) III-happen-PRS III-catch-CVB fish(III) khan-OBL-DAT
toƛ-a ***b-iž-an=ƛen***]
sell-INF **III-take-INTFUT=QUOT**
'He wants to catch the fish and sell it to the khan.'

5 In elicitation, it is also possible to use the simple present. This tense can also be used with future time reference.

I suppose that the constructions in (33a) and (33b) started out from constructions with overt CTPs that explicitly conveyed the desire that some situation may be realized in the future. Furthermore, the complement clause contained a verb in the intentional future that further reinforced the desiderative meaning and was also marked with *=ƛen*. The possibility of using the intentional future in such a context can be explained by analogy with reported speech. In quotes of reported speech constructions usually first person pronouns are used to express coreference between the author of the quote (who can be third person) and an argument or adjunct in the quote (cf. [28a] above), that is, the quote itself can have a first person subject. Therefore, in reported speech constructions the condition for the use of the intentional future can easily be met. And complement constructions as in (34) that have developed from reported speech constructions preserved this property.

In the course of time the matrix clause was left out and only what was originally a complement clause has survived; see (33a) and (33b). Similar phenomena of subordinate clauses or verb forms being used in main clauses have been called 'insubordination' or 'desubordination' in the literature (cf. Evans 2007, Jendraschek 2009). That the grammaticalization process is still in progress can be seen from the synchronic occurrence of examples such as (34) containing a CTP and examples such as (33a) and (33b) without a matrix clause. Furthermore, it is even possible to find the quotative enclitic on usual purpose clauses (35) or infinitival clauses with purposive meaning.

(35) *hadu* [*azu* *b-eta:-z=**ƛen***] *y-iƛ'i-n* *močo-ɬ-er*
3SG.F crop(III) III-mow-PURP=**QUOT** II-go-UWPST field.OBL-CONT-LAT
'She went to the field in order to mow the crop.'

There is another variant of desiderative clauses with the quotative enclitic. This variant contains the verb *-ɯɬi-* 'begin' and another verb in the intentional future marked with *=ƛen*, but the construction also means 'want'; cf. (36).

(36) *hadu* [*y-ot'-an=**ƛen***] *y-ɯɬi-ya,* *karawat-mo-ƛ'o* *xoddo*
3SG.F II-lay-INTFUT=**QUOT** II-begin-LOC.PTCP bed-OBL-SPR husband(I)
Ø-aši-yo *gom*
I-find-CVBI COP.PRS.NEG
'When she wants to sleep, she does not find her husband on the bed.'

In sum, *=ƛen* complements are formally less dependent than *-ɬi* complements and infinitive / purposive converb complements, but more dependent than zero-marked complements, because they contain a finite verb but also the quotative enclitic as a complementizer. Except for the desiderative function just described (potential type) complements marked with *=ƛen* are of the fact type or of the direct speech type. They have independent time reference and mostly realis but

occasionally also irrealis meaning. They contain verbs in the indicative mood or in non-indicative moods. The fact complements denote propositions that can be epistemically neutral or express uncertainty.

7 Summary of complementation types

The formal and semantic properties of the major Hinuq complementation types are shown in Table 2. As this table illustrates, the less formally dependent a complementation type is, the less semantically dependent it is (i.e. realized through independent time reference).

Table 2: Major complementation types and their properties

Complementation type	**Zero-marking**	***=λen***	***-ɬi***	**Infinitive/ purposive**
Form	fully independent	relatively independent	dependent	fully dependent
Preferred meaning	realis	realis (irrealis)	realis	irrealis
Time reference	independent	mostly independent	independent	dependent
Semantic type(s)	fact, activity, direct speech	fact, direct speech, potential	fact, activity	potential

8 Semantic choices in complementation

Many verbs allow for a choice between two or even more complementation types. Looking in more detail at the conditions based on which one or the other type is chosen and at minimal pairs (if available) reveals additional insights into the semantics of the various complementation types. As this paper is largely dedicated to the zero strategy, *-ɬi*, and *=λen*, I will mainly look at contexts where speakers can choose between these three types.

8.1 No marking with *-ɬi* vs. -marking with *ɬi*

Almost all verbs that allow for the *-ɬi* type of complement also allow for zero-marked complements or complements containing non-finite verb forms. In the

following, I will restrict myself to complement clauses with the matrix verb *-eq'i-* 'know' because this verb is by far more frequent than any other of the matrix verbs allowing for this kind of choice.

Usually *-ɬi* is used to mark complements as 'declarative' or 'epistemically neutral' and is thus comparable to English 'that' complements (see examples [24a]–[24f] and [25] above). Verbs that make use of *-ɬi* do not imply uncertainty about their complement proposition. In contrast, if the complement clause contains an embedded question, be it a WH-question as in (37a) or a polar question as in (37b) and (37c), then speakers commonly use the zero-marking strategy or non-finite verb forms. Similarly, if the speaker wants to express uncertainty about the proposition expressed by the complement then s/he usually employs the zero strategy or a non-finite verb form. Such examples are similar to the use of *whether* or *if* in English. This function is especially obvious in cases where the matrix verb is negated and thus the speaker explicitly states that the truth value of the proposition in the complement is unknown to him.

(37) a. [*t'ok'aw* *nido* *hago* *Ø-iƛ'i-n*] *łuzqen* *Ø-eq'i-yo* *gom*
more where 3SG.M I-go-UWPST nobody I-know-CVBI COP.PRS.NEG
'Nobody knows where he went.'

b. [*hazes* *berten* *zoq'we-ye* *zoq'we-y-me*] *dіž* *b-eq'i-me*
3PL.GEN1 wedding(III) COP-Q COP-Q-NEG 1SG.DAT III-know-NEG
'Whether there was their wedding or not, I don't know.'

c. *dіž* *r-eq'i-me* [*zek* *qema* *r-aq'e-ƛ'os*]
1SG.DAT V-know-NEG tomorrow rain(V) V-come-HAB
'I do not know whether it will rain tomorrow.'

But this is only a strong tendency, not a rule. Occasionally even in those contexts that imply uncertainty speakers may choose to use a non-finite complement clause plus *-ɬi*, (38a), or they may choose a non-finite verb form in a context that does not imply uncertainty but only lack of knowledge, (38b).

(38) a. *łuzqen* *r-eq'i-me* [*se* *r-iq-ƛ'os-**ɬi***, *taliћ=e* *balah=e*]
nobody V-know-NEG what V-happen-HAB-**ABST** happiness=Q misfortune=Q
'Nobody knows what will happen (to us), happiness or misfortune.'

b. [*hago* *šayix* *rek'we* *goɬiš=no*] *Ø-eq'i-š* *gom*
3SG.M saint man be.CVB=and I-know-RES COP.PRS.NEG
'That he is a saint (we/they) did not know.'

A similar tendency but again not a strict rule can be observed in Hinuq's closest relative Tsez. In this language, complements of verbs of knowledge and acquisition of knowledge are either marked with the suffix *-ɬi*, also a loan from Avar, as in (39c), or lack it, as in (39a) and (39b). But in contrast to Hinuq, the verb in the complement is always in its participle form and thus non-finite. As in Hinuq,

zero-marked complements predominately express uncertainty, as in (39a) and (39b), whereas *-ɬi* complements are epistemically neutral, see (39c).

Tsez

(39) a. [*žedur* *insan=ä* *elo* *ciq-es* *posu* *yaɬru*] *xoxeɬ-no*
3PL.DAT human=Q there forest-GEN1 cattle COP.PST.PTCP get.seperated-UWPST
anu
COP.PRS.NEG
'They could not distinguish whether there is a person up there or an animal.'
(Abdulaev & Abdullaev 2010: 99)

b. *nesir* [*didiw* *zaral* *žeda* *b-adi-ru=n*] *b-iy-x-anu*
3SG.M.DAT which harm 3PL.ERG III-do-PST.PTCP=and III-know-PRS-NEG
'Neither does one know what harm they did to him.'
(Abdulaev & Abdullaev 2010: 62)

c. [*ža* *nesi-s* *uži* *yaɬru-**ɬi**=n*] *Ø-iy-n*
SG.M 3SG.M.OBL-GEN1 son(I) COP.PST.PTCP-**ABST**=and I-know-UWPST
'(He) knew that he was his son.'
(Abdulaev & Abdullaev 2010)

Looking at Bezhta, another member of the Tsezic branch which, however, belongs to the East-Tsezic subgroup, we find that it does not appear to distinguish between epistemically neutral complements and complements denoting uncertainty in the same way as Hinuq and Tsez. Bezhta also makes use of *-ɬi* for marking complements. But a closer look at the Bezhta corpus at my disposal does not to indicate that the presence or absence of *-ɬi* has anything to do with certainty.

Until now we restricted ourselves to verbs of knowledge. Another group of verbs that allow but do not require their complements to be marked with *-ɬi* are perception predicates. Zero-marked complements of perception verbs or complements headed by the habitual participle are usually of the activity type; see (16a) and (16b) above, and (40). With these verbs it does not seem to be the case that the absence of *-ɬi* implies uncertainty.

(40) [*berten-i-ža-ɬ-er* *yaɬuni* *q'wiya* *kompaniya-za-ɬ-er* *hagbe*
wedding-OBL-OBL.PL-CONT-LAT or other company-OBL.PL-CONT-LAT 3PL
cadaq *b-iƛ'i-n*] *ɬuzqen* *b-ike-s-me=ƛ*
together HPL-go-UWPST nobody HPL-see-PST-NEG=NARR
'Nobody saw them going together to weddings or to other parties.'

Complements marked with *-ɬi* can be of the activity type if they contain the additional WH-word *deru* 'how' as in (41a), or they are of the fact type (41b). What happens if the matrix predicate is negated still needs to be tested (e.g. 'not see whether') since there are no such corpus examples.

(41) a. *ked-ez* *Ø-ike-s* [***deru*** *ʔali* *čeqi-do* *Ø-iƛ'i-š-**ɬi***]
girl-DAT I-see-PST **how** Ali(I) forest.in-DIRC I-go-RES-**ABST**
'The girl saw how Ali went to the forest.'
b. *ked-ez* *b-ike-s* [*uži-y* *ɣwe* *zok'-iš-**ɬi***]
girl-DAT III-see-PST boy-ERG dog(III) beat-RES-**ABST**
'The girl saw that the boy beat the dog.'

8.2 Zero-marking vs. *=ƛen*

Verbs that allow for zero-marking or the quotative enclitic *=ƛen* are the verbs of speech and the verb 'remember'. With 'remember' the difference between using the quotative enclitic and zero-marking reflects a difference between more or less formal and semantic integration (see the discussion at the end of Section 4).

When asked Hinuq speakers attribute the difference between reported speech by means of the zero strategy and reported speech by means of *=ƛen* to a distinction between direct and indirect speech. The two complementation types can be formally distinguished, see Table 3, but there are only a few differences. None of the two constructions employs non-indicative verb forms or sequence of tenses (see Sections 4 and 6 above). Zero marking is clearly less frequent, and it may be even rejected in elicitation, although it is attested in the corpus. It seems that the preference for one or the other strategy entirely depends on the speaker and probably also on the genre of text.

Table 3: Zero marking vs. quotative enclitic in reported speech

Zero marking	***=ƛen***
– no overt marking	– quotative enclitic *=ƛen*, optionally accompanied by the quotative particle *eƛin*
– the quote must occur sentence-finally after the verb of speech	– the quote can occur in initial, medial and final position
– no use of reflexive pronouns; see (42a)	– occasional use of reflexive pronouns in the quote for referents that are identical with the author of the quote; see (30) and (42b)

(42) a. *Murad* *iyo-qo* *hardezi* *Ø-iqqo* *"de* / **zo* *idudo* *Ø-iž-o!"*
Murad(I) mother-AT request I-happen.PRS 1SG / REFL home I-take-IMP
'Murad begs his mother, "Take me home!"'
b. *Murad* *iyo-qo* *hardezi* *Ø-iqqo* *"de* / *zo* *idudo* *Ø-iž-o=ƛen!*
Murad(I) mother-AT request I-happen.PRS 1SG / REFL home I-take-IMP=QUOT
'Murad begs his mother, "Take me home!"'

8.3 *-ɬi* vs. *=ƛen*

There are only very few predicates that allow both for the *-ɬi* and for the *=ƛen* type of complements. These verbs are basically 'remember' and predicates of knowledge and acquisition of knowledge (except from *-eq'i-* 'know'). The semantic distinction between these two complementation types is as follows: *=ƛen*-marked complements express propositions the knowledge about which has been acquired from oral report. Although this is not directly stated in (43a) because the matrix verb is not a verb of speech, it can be implied from the use of the quotative enclitic. If a follow-up clause contains a statement that contradicts this implication, the whole sentence is unacceptable. Whenever *-ɬi* marks the complement clause the source of the information is unclear and unimportant. In such cases the proposition in the complement clause is merely presented as a fact; see (43b).

(43) a. *dіž* *c'aɬ-iš* [*Musa* *Ø-aq'e-s=**ƛen***], *(*dіž* *hayło-s*
1SG.DAT get.to.know-PST Musa(I) I-come-PST=**QUOT** 1SG.DAT 3SG.M.OBL-GEN1
ƛ'oq'on *b-ikey-ƛo)*
hat(III) III-see-SML
'I got to know that Musa came (*when I saw his hat).'

b. *de* *hayło-qo* [*me* *aƛ-a-do* *Ø-aq'e-ƛ'os-**ɬi***]
1SG.ERG 3SG.M.OBL-AT 2SG village-in-DIRC I-come-HAB-**ABST**
c'aɬ-er-an
get.to.know-CAUS-INTFUT
'I let him know that you will come to the village.'

In complement clauses to which *=ƛen* has been attached it is possible to use non-indicative moods such as the imperative even if the matrix verb is not a verb of speech. In (44) the use of *=ƛen* implies that the sentence depicts a speech act during which the speaker gave a command to a group of people.

(44) *de* *haze-qo-r* *r-eq'i-r-iš* [*zek* *šaharli-do* ***nox=ƛen***]
1SG.ERG 3PL.OBL-AT-LAT V-know-CAUS-PST tomorrow town.in-DIRC **come.IMP=QUOT**
'I told them (lit. made them know) to come to the town tomorrow.'

8.4 Other alternative choices

There is a fourth interesting choice between complementation types that is worth to look at, namely the choice between a case-marked participle or masdar and the quotative enclitic *=ƛen*. This choice is available with the propositional attitude predicates 'believe', 'doubt', 'be sure', 'think', 'wonder' and the verbs 'be happy, agree', 'be disappointed' and reported speech. Employing case-marked

participles/masdars results in complements of the activity type, see (45a), while using the quotative enclitic gives complements of the fact type, see (45b).

(45) a. *di* *rok'we* *r-aq'-o* [*obu* ***Ø-aq'-a-nu-ƛ'o***]
1SG.GEN1 heart(V) V-come-PRS father(I) **I-come-INF-MSD-SPR**
'I am sure about the coming of my father.'

b. *di* *rok'we* *r-aq'-o* [*zek* *di* *essu* *Ø-aq'e=**ƛen***]
1SG.GEN1 heart(V) V-come-PRS tomorrow 1SG.GEN1 brother(I) I-come=**QUOT**
'I am sure that my brother will come tomorrow.'

The fact that *=ƛen* complements are of the fact type and refer to utterances rather than to real situations can even be made more explicit through adding the quotative particle *eƛin* to the clause. Thus, in (46a) the mullah agrees and expresses his happiness by saying *noxan* 'I will come.', that is, the complement clause represents a direct quote of the mullah's utterance. Compare this example with (46b) which contains the same matrix predicate *razi -iq-* 'be happy, agree'; here it is the event about which the subject Patimat is happy.

(46) a. *hago* *ʔoloqanaw* *dibir* *razi* *Ø-iq-no* *"nox-an"=**ƛen*** ***eƛin***
this young mullah agree I-happen-UWPST come-INTFUT=**QUOT** **say-CVB**
'The young mullah agreed that he would come.'

b. *Pat'imat* *razi* *y-iqqo* [*iyo-y* *buƛe* ***b-ux-oru-ƛ'o***]
Patimat happy II-happen.PRS mother-ERG house(III) **III-buy-PST.PTCP-SPR**
'Patimat is happy about her mother buying a house.'

As mentioned in Section 3, a masdar to which the first genitive suffix is added can head a complement that refers to the topic of a conversation. Despite the fact that in this case the whole sentence is about a speech event, the complement nevertheless directly refers to the event itself and belongs thus to the activity type; see (47a). Only if the complement contains *=ƛen*, it is referring to the assertion (fact type) and not to the event itself; see (47b).

(47) a. *Maħama-y* *xabar* *b-uː-s* [*zoni* *aqili*
Mahama-ERG story(III) III-do-PST REFL.SG.ERG woman(II)
y-iž-a-nu-s]
II-take-INF-MSD-GEN1
'Mahama talked about taking a wife' (i.e. marrying).

b. *hayłoy* *xabar* *kekir-no* *"zonzo* *omoq'i-ya-y* *mecxer*
3SG.M.ERG story send-UWPST REFL.SG.GEN2 donkey-OBL-ERG money
*gotto"=**ƛen***
pour.PRS=**QUOT**
'He said that his donkey poured money.'

The semantic distinction goes hand in hand with the formal distinction: case-marked participles/masdars have more noun-like properties (e.g. no TAM suf-

fixes, case suffixes; past participles occur in headless relative clauses filling argument positions in the main clause) than complement clauses with *=ƛen* (finite predicates, full TAM paradigm, including non-indicative mood), although both complementation types fully preserve the argument structure of the embedded verb (e.g. A argument in the ergative case and O argument in the absolutive case if the embedded verb is transitive, cf. [47a] and [47b]).

There is a fifth alternation in the marking of complement clauses of the verb 'remember' that allows for fact-type complements to be marked with *-ɬi* or *=ƛen* and for potential-type complements to be marked with the infinitive/purposive converb (see Section 4).

9 Combination of complementation types

The different complementation types are normally not combined within one and the same complement construction (though iterative embedding is possible). The only exception is the sporadic combination of infinitive/purposive converb with the quotative enclitic *=ƛen*. However, examples such as (48) are rare; normally speakers choose only one of the two strategies. Yet Hinuq is not the only Nakh-Daghestanian language that combines the infinitive and the quotative marker for expressing purpose clauses; see Section 12.1 for other languages.

(48) *xozyayn* [***Ø-ez-a=ƛen***] *Ø-aq'e-yƛ'o,* *b-ik-o* *gome* *hayłoz*
owner(I) **I-look-INF=QUOT** I-come-SML III-see-CVBI COP.PRS.NEG 3SG.M.DAT
hayłiteł *boc'e*
there wolf(III)
'When the owner came to take a look he finds the wolf there, right?'

At this point it is appropriate to add a few words on the influence of Russian. A number of languages spoken on the territory of Russia or in areas with a long influence from Russian have borrowed Russian complementizers, notably the complementizer *čto* 'what, that' (cf. Kehayov, this volume; Klumpp, this volume). Hinuq has borrowed an increasing number of Russian words, but until now shows only very weak Russian influence in the syntax. Thus, young speakers may rarely use Russian conjunctions such as *esli* 'if' or *i* 'and' for the formation of adverbial clauses or in clause coordination, but they always combine them with the native Hinuq means of forming such clauses (e.g. if *esli* appears in an adverbial clause the head of the clause has necessarily the conditional converb suffix attached to it). Similarly, young Hinuq speakers very rarely employ Russian *čto* in complement clauses, especially when they translate from Russian. But in any case *čto*

must be combined with the appropriate Hinuq strategy of forming a complement clause, e.g. with the quotative enclitic; cf. (49).

(49) *hoboži* *hałо* *rek'u-z* *bič'i* *r-iq-no* [***čto*** *čiyo*
now this.OBL man-DAT understanding(V) V-happen-UWPST **that** salt(V)
ič'č'a *r-egi* *goł=**ƛen***] *eƛi-n*
very V-good COP.PRS=**QUOT** say-CVB
'Now the man understood that salt is the best thing.'

10 Other types of matrix predicates

Occasionally the complement-taking element is not a verb but an adjective, an adverb or a noun. Adjectives or adverbs can take *-łi* complements or infinitive clauses; see (50) below and (24e) above. Note that in (50) the word *behula* is an Avar loan stemming from the simple present form of the Avar verb *b/eh/ize* 'be possible, be probable'.[6] This construction does not require a Hinuq finite verb.

(50) [*hago* *zek* ***Ø-aq'-a***] *behula*
3SG.M tomorrow **I-come-INF** possible
'It is possible that he will come tomorrow.' or 'It is possible for him to come tomorrow.'

Complements of nouns have various forms. They can have the same structure as complements of verbs; that is, they can be headed by an infinitive/purposive converb, as in (50) and (51a), or by a finite structure marked with the quotative enclitic, as in (51b). Another possibility is to have a participle, which is also used for the formation of relative clauses, but where the complement-taking noun cannot be regarded as the head of a relative clause since it does not fulfill a grammatical role in the participial clause; see (51c). Such constructions are rather common in Nakh-Daghestanian languages (see Section 12.2 below). The noun is not necessarily adjacent to its complement, and it can precede or follow it.

(51) a. [*ƛ'ere-do* *dew-qo* ***moł-ayaz***] *diž=no* *ʔilmu* *y-eq'i-yo*
up-DIRC 2SG.OBL-AT **teach-PURP** 1SG.DAT=and science(IV) IV-know-CVBI
gom
COP.PRS.NEG
'I do not know more science to teach you.'

6 Hinuq has also borrowed the Avar adjective *behulew* 'possible', which occurs in adjectival functions (attributive, predicative). In contrast, *behula* cannot fulfill adjectival functions in Hinuq, but appears only in complement clauses.

b. *dіž* *toq-iš* *xabar* [*xan* *Ø-aq'e-n=**ƛen***]
1SG.DAT hear-PST story khan(I) I-come-UWPST=**QUOT**
'I heard the news that the khan came.'

c. *dіž* [*xan* ***Ø-aq'o*** ***goła***] *xabar* *b-eq'i-yo*
1SG.DAT khan(I) **I-come-CVBI** **COP.PTCP** story(III) III-know-PRS
'I know the news that the khan came.'

11 Complementation markers in adverbial and relative clauses

As occasionally mentioned throughout this paper, some of the markers occurring in relative clauses are also used in other types of subordinate clauses. In the first place, this is the narrative converb, the copula converb, and the habitual participle. The function of the narrative converb in complement clauses of the activity types was illustrated in (10). However, this is just a marginal use of the narrative converb which is very frequently used to coordinate two or more simultaneous or sequential events. In this function the narrative converb expresses anteriority or perfectivity, and the respective clauses can normally be translated by using the coordination 'and'; cf. (52a). Mainly with verbs of motion, but also with other verbs the narrative converb also expresses manner of action; cf. (52b).

(52) a. [*gołіš* *mecxer=no* ***b-iži-n***] [*mesed=no* ***b-iži-n***]
COP.CVB money(III)=and **III-take-CVB** gold(III)=and **III-take-CVB**
[*ukru=n* ***b-iži-n***] *Ø-iƛ'i-n*
silver(III)=and **III-take-CVB** I-go-UWPST
'(He) took his money, his gold and his silver and went (there).'

b. ***qaƛe-n*** *Ø-aƛ'i!*
shout-CVB i-talk.IMP
'Talk louder!'

The copula converb *gołіš* is used with an adpositional meaning 'being, having, with'. It heads what seems to be a kind of adverbial clause, e.g. *hunar gołіš rek'we* (lit. talent COP.CVB man) 'a man with talent'; see also (52a) above. The habitual participle is mostly used to modify nouns in short relative clauses with an adjectival character that describe characteristic properties of the head noun. It expresses either relative future time reference of the relative clause or characteristic properties of the head noun with no specific time reference.

(53) *de* *es-an* [*dіž* ***b-eq'i-ƛ'os***] *xabar*
1SG.ERG tell-INTFUT 1SG.DAT **III-know-HAB** story(III)
'I will tell the things that I know.'

As has been shown above, the past participle and the masdar must be case-marked in order to occur in complement clauses; cf. (5b) for the past participle and (6), (7), (45a), (47a) for the masdar. Without the case suffixes or with different case suffixes the past participle and the masdar are also found in relative or adverbial clauses. The past participle occurs in headed and headless relative clauses with past time reference or with a kind of perfect meaning; see (54).

(54) [***t'ot'er-omeru***] *rek'u-z* *ni=n=tow* *r-eg* *gom*
learn-PTCP.PST.NEG man.OBL-DAT where=and=EMPH V-well COP.PRS.NEG
'The man who did not study lives badly everywhere.'

The participle also heads adverbial clauses conveying the meaning of immediate anteriority 'as soon as'; cf. (55). When it is case-marked with the first genitive suffix or certain spatial case markers, other types of adverbial clauses are created (e.g. comparative constructions with the meaning 'the X-er the Y-er' and clauses expressing causes). More information on the Hinuq past participle can be found in Forker (2013: 255–257, 259–262, 570).

(55) *hayi-š* [*hag* ***eƛi-yoru***] [*r-iy-no* *ižey*] *toƛ-o* *gaqinnuk'a*
there-ABL.1 that **say-PTCP.PST** V-take-CVB eye(V) give-PRS bad
essu-z
brother-DAT
'As soon as he said that, he pulled his eye out and gave it to the bad brother.'

The occurrence of the masdar in relative clauses is rather rare. The masdar in relative clauses has a prospective meaning and refers to events that have not happened yet but are strongly expected to happen in the future:

(56) *hadu* goł [*tetu* ***y-iy-a-nu***] *cen*
this COP.PRS cream(IV) **IV-take-INF-MSD** brynza
'This is cottage chese from which the cream has not been taken away.'

The abstract suffix is used to form a converb that occurs in adverbial clauses with the temporal meaning of simultaneity. The relavant example (22) can be found in Section 5. The other complementation markers (zero strategy, infinitive/purposive converb, quotative enclitic, and clause union) are not found in adverbial or relative clauses.

12 Some notes on complementation types in Nakh-Daghestanian languages

In this section I will briefly provide some information about complementation types and their semantics in Nakh-Daghestanian languages. This allows a comparison with the Hinuq data and the embedding of this data in the wider context of Nakh-Daghestanian languages. I will first present a sketch of the major and some minor complementation types, and then discuss syntactic and semantic variations between different types and the possibility for their combination.

The major complementation strategies in Nakh-Daghestanian languages are:

1. non-finite verb forms (infinitive, masdar, participles, converbs)
2. quotative particles or enclitics
3. zero marking

Minor strategies are:

1. complementizers
2. subjunctive mood
3. relative clause-like constructions

The major strategies can probably be found in (almost) all languages of the family, whereas the minor strategies are attested only in a few languages. Complementation types familiar from European languages such as complementizers and subjunctive mood are quite untypical for Nakh-Daghestanian languages.

12.1 Major complementation strategies in Nakh-Daghestanian

All Nakh-Daghestanian languages have infinitives or verbal forms similar to infinitives occurring in complement clauses with equi-deletion (i.e. the subject of the complement clause is coreferential with an argument in the matrix clause and is therefore left unexpressed in the complement clause). In Tsakhur, it is the 'potentialis' that has similar function to a standard infinitive, and in Icari Dargwa it is the 'subjunctive', a verbal form historically related to infinitive forms in other Dargi languages, but showing full-fledged personal agreement; cf. (57). Typical CTPs that make use of the infinitive are, for example, modal verbs, the phasal verb 'begin' (but not 'stop, finish'), 'want', 'teach', 'help'.

Icari Dargwa

(57) [*ca* *ka<b>ič-ib* *Xabar* ***b-urs-ij***] *b-ik:-ul-da*
one happen<N>.PFV-PRET story(N) **N-tell.PFV-SBJV** N-want.IPFV-CVB-1
'I want to tell a true (lit. 'happened') story.'
(Sumbatova & Mutalov 2003: 108)

A common complementation type found in Nakh-Daghestanian languages is the masdar. Masdar complements usually belong to the activity type which is especially obvious if they are compared with alternative complementation types (see Section 12.3 below). The masdar has nominal and verbal properties. It gets its case assigned by the matrix verb or, rarely, by the construction as a whole. This means, the masdar can be in the absolutive case or in another case, depending on the valency frame of the CTP. In many Nakh-Daghestanian languages verbs such as 'believe', 'be happy with, agree', 'be glad' and 'wonder' or 'doubt' take masdars marked with spatial cases as their complements, see (58a), but other matrix verbs assign other cases, as in (58b). It seems that in most languages the masdar – when functioning as the head of a complement clause – keeps the valency frame of the verb, that is, it assigns case to its arguments exactly in the same way as a finite verb would do.

Archi, Chechen

(58) a. *zon* *x̌ʷara* [*un* ***w-aL-mul-li-tti-š***]
1SG be.glad 2SG **M-come-MSD-OBL-SPR-ABL**
'I am glad about you (masc.) coming.'
(Kibrik 1994: 357)

b. *Mu:sa:s* *xaʔi:tira* [*ša:* *üöğaz-vaxã* ***xilar***]
Musa.ERG know.CAUS.PST REFL offend-POT **be.MSD**
'Musa let it be known that he was offended.'
(Nichols 1994: 62)

In Khwarshi, Tsakhur and Bagvalal it is also possible to have the masdar assign the genitive case to one argument of the embedded clause. This is usually the S or A argument, but rarely also an argument bearing another grammatical role such as the O or the experiencer. In Bagvalal, the choice between the genitive marking of the masdar's most prominent argument and the preservation of the normal valency frame depends on the referentiality of the complement clause (see Section 12.3 below). In the other languages it seems that the meaning of the two variants is identical: e.g. in (59a) from Tsakhur the masdar assigns the expected cases to its arguments (AD-elative to the experiencer and absolutive to the stimulus), in (59b) the experiencer has the attributive marking (roughly iden-

tical to a genitive), and in (59c) the stimulus has the attributive marking. Note that at most one argument can take the attributive case.[7]

Tsakhur

(59) *bajram-i-s* *q'abił-e:x-e* *wo-d*
Bajram-OBL-DAT like-be.II-IPFV COP-IV
(Ljutikova & Bonč-Osmolovskaja 1999: 491)
a. [*dak-i-ss-e* *Pa:t'imat* *k'ele<r>Xin-i:*]
father-OBL-AD-ELA Patimat(II) forget<II>.PFV-MSD(IV)
(Ljutikova & Bonč-Osmolovskaja 1999: 491)
b. [*dakk-i-n* *Pa:t'imat* *k'ele<r>Xin-i:*]
father-OBL-ATTR Patimat(II) forget<II>.PFV-MSD(IV)
(Ljutikova & Bonč-Osmolovskaja 1999: 491)
c. [*dak-i-ss-e* *Pa:t'imat-i-n* *k'ele<r>Xin-i:*]
father-OBL-AD-ELA Patimat(II)-OBL-ATTR forget<II>.PFV-MSD(IV)
(Ljutikova & Bonč-Osmolovskaja 1999: 491)
'Bayram likes that father forgot Patimat.'

Furthermore, in Hinuq, Khwarshi and Bagvalal the topic of a conversation or of an act of cognition can be expressed by using a masdar in the Genitive (cf. [47a] above).

In some languages (Hinuq, Khwarshi, Bagvalal) the masdar alternates with participles, that is, there are verbs that allow for participle and masdar complements. In contrast to Hinuq, in Khwarshi and Bagvalal the formal difference correlates with the semantic difference between activity type (masdar) and fact type (participle) (see Section 12.3). The following examples come from Khwarshi, whereby (60a) illustrates the activity type and (60b) the fact type. The Khwarshi nominalizing suffix *-łar* in (60b), which can optionally be added to the participle, is a cognate of Hinuq *-łi* (Section 5).

Khwarshi

(60) a. *dil* *l-iyo:q'* [*iso* *l^{j}-uƛƛo* *kеč'i* ***b-ez-nu*** *bani-ma*]
1SG.LAT IV-know.GNT 3SG.GEN.1 IV-strong song(III) **III-take-MSD** shower-in
'I know about his singing loudly in the shower.'
(Khalilova 2009: 367)

7 If the standard valency frame is preserved, long distance agreement is possible as alternative to local agreement, i.e. in (59a) the auxiliary would be *wo-r* 'COP-II' instead of *wo-d* 'COP-IV'. Sentences such as (59b) and (59c) admit only local agreement. Note also that in example (59c) the embedded argument in the attributive (genitive), *Pa:t'imat-i-n*, triggers agreement on the masdar. This is unusual since normally only absolutive arguments can trigger gender/number agreement.

b. *dil* *l-iyo:q'* [*iso* *l^{i}-uƛƛo* *kеč'i* ***b-ez-dow-(ɬar)***
1SG.LAT IV-know.GNT 3SG.GEN1 IV-strong song(III) **III-take-GNT.PTCP-NMLZ**
bani-ma]
shower-in
'I know that he sings loudly in the shower.'
(Khalilova 2009: 367–368)

Matrix predicates that allow for the participle strategy are typically 'know', perception verbs such as 'see' and 'hear', and other verbs denoting cognitive activities, like 'forget' or 'remember'. In Avar the situation is similar to Hinuq (recall Section 8.1): complement clauses of 'know' are headed by a participle if they contain an indirect question (61); otherwise they are headed by a finite verb or participle plus the nominalizing suffix *-ɬi*, as in (19) above.

Avar

(61) *dida* *ɬa-la-ro* [***kin*** *duca* *q'abul* ***hab-ile-b*** *dun*]
1SG.LOC know-PRS-NEG **how** 2SG.ERG accepting **do-PTCP.PRS-III** 1SG
'I do not know how you will receive me.'
(Alekseev & Ataev 1998: 111)

The use of converbs is another relatively common complementation type. It predominantly occurs with phasal verbs (Ingush, Khwarshi, Hinuq, Bagvalal, Godoberi, Icari Dargwa) and perception verbs (Lezgian, Bagvalal, Godoberi, Avar) and occasionally with 'want, like, be pleased' (Ingush, Chechen, Lezgian, Icari Dargwa). The converbs occurring in such constructions often have a perfective, past or anterior meaning; see (62).

Ingush

(62) *shi* *shu* *dealcha* *so* [***diishaa***] *jistie+joal*
two year D.go.CVBTEMP 1SG **D.study.CVBANT** J.finish+J.AUX.PRS
'In two years I finish my studies.'
(Nichols 2011: 555)

The vast majority of Nakh-Daghestanian languages have quotative particles, either realized as enclitics or as independent phonological words, which usually have been grammaticalized from a converb form of the verb 'say'. These particles mark reported speech in all languages which have them; cf. (63a). Matrix verbs are not only verbs such as 'say', 'tell' or 'ask', but also other verbs that denote actions involving speech, e.g. 'swear', 'command', negate', etc. They also mark complements of some propositional attitude verbs (e.g. 'think', 'believe', 'hope', 'fear', 'regret', 'suppose'), as in (63b), and of some other verbs such as 'seem', 'remember' or 'hear'. The particles occur at the end of their complement clauses.

In these constructions the embedded verbs are finite, and there is usually no deictic change apart from the use of special reflexive or logophoric pronouns (see the discussion about Hinuq in Sections 6 and 8.2). Furthermore, in some languages (Ingush, Godoberi, Hinuq, Tsez, probably also Tsakhur) the quotative particle can be optionally added to the infinitive in order to form a purpose clause; cf. (63c). The following examples come from Godoberi where the quotative particle *-L'u* is a grammaticalized form of the converb *hiL'u* 'having said'. Often the vowel preceding *-L'u* is lengthened, reflecting the loss of the syllable *hi-*.

Godoberi

(63) a. *jaš-u-di* *hiL'i* *il-u-qi* [*den* *quča*
daughter-OBL-ERG say.PST mother-OBL-AD 1SG.ERG book(N)
*b-aX-i-Li-bu-da-**L'u***]
N-bring-INF-FUT-PTCP-COP-**QUOT**
'Daughter told mother: "I will bring the book."'
(Haspelmath 1996: 182)

b. *waša* *urR-ant-a-da* [*X^{w}aji* *hawa-waqi* *ida-**L'u***]
boy think-PRS-CVB-COP dog here-PRT COP-**QUOT**
'The boy thinks that the dog is here.'
(Haspelmath 1996: 185)

c. *muradi* *wu-na* *buɬiXa* [*kofe* *b-aX-i* / *b-aX-i:-**L'u***]
Murad M-go.PST to.Botlikh coffee(N) N-take-INF / N-take-INF-**QUOT**
'Murad went to Botlikh to buy coffee.'
(Haspelmath 1996: 185)

Zero-marking is also found in a substantial number of Nakh-Daghestanian languages (Tsova-Tush, Ingush, Chechen, Khwarshi, Lezgian, Archi, Tsakhur, Avar, and probably more). In general it alternates with the use of the quotative particle because in the most languages that allow for zero-marking reported speech is one of the contexts in which this complementation type can be found. The second context where zero-marking is often attested is complements of the verb 'know', 'let know', 'seem', and a few other verbs.

12.2 Minor complementation strategies in Nakh-Daghestanian

Minor complementation strategies are those that are found only in some languages of the family. For instance, the use of complementizers is not very common. A few Lezgic languages and Khinalug have borrowed canonical complementizers ultimately going back to the Persian *ke*, which entered the languages via Turkic: Lezgian *xîi*, Tsakhur, Kryz and Khinalug *ki* (cf. Authier 2009: 289–290 for Kryz and Daniel 2013 for Khinalug). These complementizers mainly occur in

reported speech, and in contrast to the native quotative particles, they introduce a finite complement clause. Apparently they are not combined with any other native complementation marker.

Only a few languages have native canonical complementizers. Tsova-Tush has the complementizer *me* 'that' used with matrix verbs such as 'say', 'try', 'want', verbs of knowledge and cognition. It can be combined with the subjunctive mood (see Section 12.3). Tsakhur has the complementizer *naXa* 'if, whether' that can be combined with the quotative particle or with the masdar (see Section 12.4). It is never employed for indirect embedded questions, but occurs in complements of 'think', 'fear' and with negated matrix predicates such as 'know', 'hear', 'see', 'believe' or 'remember'. It behaves similarly to focus particles because it is encliticized to the focused element of the embedded clause. Finally, Ingush has two conjunctions *sanna, muo* 'like, as' occurring in complements of verbs such as 'dream', 'think', and 'seem'.

Only the three Nakh languages Tsova-Tush, Chechen and Ingush have complement clauses making use of subjunctive verb forms. These verb forms, which are built from various finite or non-finite verb forms by adding a suffix to them (Chechen *-jla*, Ingush *-aljga*, Tsova-Tush *-lo*) are almost exclusively used in complement clauses. In Tsova-Tush, the subjunctive is also used in purpose clauses; and it can be combined with the neutral complementizer *me* 'that' (see Section 12.3). The following example is from Chechen where the verbs 'want' and 'know' can take subjunctive complements, but also indicative complements with apparently no change in meaning:

Chechen

(64) *süöga* *Mu:sa:s* *xaʔi:tira* [*šiĕ* *da:* *velli:la* / *velli*
1SG.ALL Musa.ERG know.CAUS.PST REFL.GEN father(V) V.died.SBJV / V.died
'Musa let me know that his father has died.'
(Nichols 1994: 63)

A number of languages (Lezgian, Tsakhur, Bagvalal, and Avar) have constructions that are formally relative clauses (i.e. participial clauses) but semantically they are similar to complement constructions. These clauses may or may not have a 'semantically empty' head noun such as 'time', 'fact' or 'thing'. In general, relative clause-like constructions that are formed by means of participles but where the head noun does not fulfill any grammatical role in the complement clause are rather common in Nakh-Daghestanian languages (Daniel & Lander 2010).[8] Thus,

8 In fact, Nakh-Daghestanian languages pose only very few, if any, restrictions on the grammatical roles or positions that can be relativized (van den Berg 2005).

it comes as no surprise that these constructions can resemble complements; cf. (65)

Lezgian

(65) *za-z* [*Zerli* *zi* *lap* *muq'wa-di* *tir*] *č'al* *či-da-j*
1SG-DAT Zerli 1SG.GEN very related-NMLZ.SG be.PTCP word know-FUT-PST
'I knew that Zerli was closely related to me.' (Lit. I know the word that ...)
(Haspelmath 1993: 373)

12.3 Choices between complementation types

Most languages permit at least some choices between different complementation types. Whether such choices are available or not depends in the first place on the matrix predicates. Most choices correlate with syntactic or semantic distinctions or restrictions, but occasionally there may not be any semantic or syntactic differences expressed by the different variants (cf. the Chechen example in [64]).

In terms of syntactic conditions regulating choices between complementation strategies, I found two types: (i) coreference constraints (usually same-subject vs. disjoint-subject) and (ii) constraints on the valency of the embedded verb and/or its grammatical roles. Thus, a number of Nakh-Daghestanian languages require different complementation types depending on whether the embedded subject is coreferent with an argument (mostly also the subject) of the matrix clause or not. For instance, in Ingush the verb 'know' can take infinitival complements if the experiencer of 'know' and the subject of the complement clause are coreferent, see (66a); otherwise it takes finite verbs as in (66b), subjunctives as in (66c), or nominalized participles.

Ingush

(66) a. *Waishietaa* [*t'ehwa-jysa*] *xou*
Aisha.DAT late-J.stay.INF know
'Aisha is good at being late.' (Lit. 'Aisha knows how to be late.')
(Nichols 2011: 553)

b. *Waishietaa* *xou,* [*Muusaa* *t'ehwa* *hana* *vysaav*]
Aisha.DAT know Musa behind why v.stay.UWPST.V
'Aisha knows why Musa was late.'
(Nichols 2011: 543)

c. *suona* [*yz* *dika* *sag* *voljga*] *xou*
1SG.DAT 3SG good person v.be.SBJV know.PRS
'I know (that) he is a good person.'
(Nichols 2011: 547)

Dependence on the valency frame of the grammatical roles can be illustrated with the following example from Bagvalal. In this language, complement clauses headed by a masdar can either preserve their valency frame, as in (67a), or assign the genitive to the subject of the embedded verb, as in (67b). The latter variant is not possible with so-called dative and affective verbs that have an experiencer in the dative or affective case and a stimulus in the absolutive (67a).[9] Similarly, Hinuq has some constraints concerning complement clauses of modal verbs such as 'must' (Forker 2013: 621–626).

Bagvalal

(67) a. [*diba* / **di-b* *čʼalʕa-n*] *men* *tʼober-eː-jšʔ*
1SG.AFFC / 1SG.GEN bore-MSD 2SG.ERG fulfill-CAUS-Q
'Did you fulfill that what bores me?'
(Kalinina 2001: 526)

b. *diha* *čʼalʕaː* [*dub* *uhi-ahi* *Xʷadiri-r*]
1SG.DAT bore 2SG.GEN here-there go-MSD
'You going here and there bores me.'
(Kalinina 2001: 525)

The following semantic distinctions expressed by alternation of complementation types can be found in Nakh-Daghestanian languages:
- fact type vs. activity type
- fact type vs. direct speech type
- realis-factive vs. irrealis-potential (e.g. *remember that* vs. *remember to*)
- embedded interrogative vs. embedded declarative
- factive vs. modal.

The fact type vs. activity type distinction was already described for Hinuq, but it is also found in other languages of the family (e.g. Khwarshi, Bagvalal, Lezgian, Tsakhur). Broadly speaking, the activity type reading is available with complements headed by a masdar, whereas the fact type reading is possible with complements containing (nominalized) participles, but there are exceptions. Table 4 summarizes the formal expression of this distinction in a number of Nakh-Daghestanian languages including Hinuq.

9 The only exceptions to this rule are the dative verb 'want' and the affective verb 'can'.

Table 4: Fact vs. activity complementation

	Fact type ('that')	Activity type ('how')
Khwarshi	nominalized participle	masdar
Lezgian	nominalized participle	general converbs
Bagvalal	participle	masdar
Tsakhur	zero marking, quotative particle *wi*, masdar	masdar, zero strategy
Hinuq	zero marking, quotative enclitic *=ǟen*, abstract suffix *-łi*, copula converb	zero marking, various non-finite verb forms (converbs, participles, masdar)

Whether such a distinction is available at all depends in the first place on the matrix predicate, e.g. in Bagvalal the semantic distinction is clearly related to the use of participles vs. masdars, but due to their semantics phasal verbs do not express the distinction. Typical matrix predicates that are able to express an activity type vs. fact type distinction are perception verbs and cognitive verbs such as 'remember', 'forget', 'be pleased' or verbs of speech. Examples (68a) and (68b) show a minimal pair from Lezgian, where the distinction is restricted to the verb 'see' (see also the Khwarshi examples in [60a] and [60b] above). In Tsakhur the masdar can be used with fact-type complements and with activity-type complements. The examples in the grammar suggest that for activity complements the masdar must be case-marked as in (68c), whereas for fact complements the case marking is optional; cf. (68d) and (59a)–(59c) above. In fact, it seems that fact-type complements of the masdar are usually not case-marked.

Lezgian, Tsakhur

(68) a. *ada-z* [*zun* *čpi-z* ***klig-zawa-j-di***] *aku-na*
he-DAT 1SG REFL.PL-DAT **look-IPFV-PTCP-NMLZ** see-AOR
'He saw that I was looking at them.'
(Haspelmath 1993: 365)

b. *ada-z* [*wiči-n* *juldaš* *ğürčeq^{h}an-ri* ***q'u-nwa-z***] *aku-na*
he-DAT REFL-GEN comrade hunter-PL.ERG **catch-PRF-CVB** see-AOR
'He saw how the hunters had caught his comrade.'
(Haspelmath 1993: 370)

c. *iči:* *šad-ejx-e* *wo-r* [*ginej* ***qež^{w}-i:-l-e***]
girl(II) happy-become.II-IPFV COP-II bread(IV) **bake.IV.PFV-MSD-SPR-ELA**
'The girl is happy about baking bread.'
(Ljutikova & Bonč-Osmolovskaja 1999: 490)

d. *bajram* *mat-e:x-e* *wo-r* [*jed-e:* *čož*
Bajram(I) wonder-become.I-IPFV COP-I mother-ERG brother(I)
doXtur-u-k'le ***hag^{w}-i:-l-e***]
doctor-OBL-AFFC **show.PFV-MSD-SPR-ELA**
'Bajram wondered that mother showed the brother to the doctor.'
(Ljutikova & Bonč-Osmolovskaja 1999: 491)

The distinction between fact type and direct speech or speech act type is attested in most if not all Nakh-Daghestanian languages. It boils down to the fact that these languages have reported speech constructions not showing any type of syntactic subordination at all. In these constructions the quote is just a repetition of what the original speaker said, containing the same pronouns, verbs forms, temporal or spatial references as the original utterance. In addition, it is possible to have reported speech constructions with reflexive pronouns and occasionally deictic expression with shifted reference. In that case complements are of the fact type. Thus, compare (69a) with (69b). Both quotes contain finite verb forms. The only difference is the reflexive pronoun in (69a) that refers to the author of the quote. In (69b) a normal first person pronoun has been employed for referring to the author of the quote. Hinuq reported speech constructions were described in Sections 6 and 8.2

Archi[10]

(69) a. *“ha”* *ja-r* *laha* *bo-li,* *“inž* *os* *dija-s:=u* *to-r*
well this-II girl.ERG say.PFV-EVID REFL one father-DAT=and that-II
biši *buwa-s:=u* *d-ak:u-s* *e<r>q’i-s* *d-a-su”* *bo-li*
alien mother-DAT=and II-see-INF arrive<II>-INF II-do-IMP say.PFV-EVID
“noL’-a-ši *os* *t’i:nna* *č’em-na”*
house-in-ALL one a.little time-in
‘“Well”, said the girl, “take me home for some time such that my father and my stepmother see me.”’

b. *“zon* *išik* *w-i”* *bo-li*
1SG here I-COP say.PFV-EVID
‘“I am here.” he said.’

The realis-factive vs. irrealis-potential distinction is fairly widespread. Usually the infinitive is used for encoding irrealis-potential complement clauses (e.g. *remember to*) and another strategy is employed for realis-factive complements (e.g. *remember that*). For example, in Tsakhur realis complements are expressed via the use of a masdar or the quotative particle *wi*, and irrealis complements via the potentialis form which corresponds to the infinitive; see (70). Thus, compare examples (59a-c) and (74) with example (70).

10 These two examples are from the Archi online corpus that can be found at http://www.philol.msu.ru/~languedoc/rus/archi/corpus.php.

Tsakhur

(70) *ičï:-s* *q'abił.ex-e* *wo-d* [*jiǯ-e:* *ginej* ***qeǯ-es***]
girl-DAT like(IV)-IPFV COP-IV REFL(II)-ERG bread(IV) **bake(IV)-POT**
'The girl likes to bake the bread herself.'
(Ljutikova & Bonč-Osmolovskaja 1999: 532)

A somewhat similar distinction is found in Bagvalal. This language has two infinitives, the 'potential infinitive' and the 'imperfective infinitive' that both occur in complement clauses. The former is far more frequently used, for example in complements of modal verbs, but also with verbs of speech or 'forget' and 'fear'. The latter is basically restricted to complements of 'force, cause', 'begin', 'know', 'learn' and 'teach'. According to Kalinina (2001), the difference can be explained by the fact that the former infinitive has a potential (irrealis-like) meaning, referring to the intentions of the subject and thus to situations that have not taken place yet. This verb form also has a hortative function (e.g. *let's do X*), which is semantically similar to the potential meaning it has in complements. The imperfective infinitive, on the contrary, when used in complement clauses denotes situations that are somewhat more real. For instance, complements of the verb 'begin' denote the initial phase of real situations. Similarly, complements of 'force' or 'teach' refer not to potential events, but to real events.

Another often attested semantic distinction is the difference between embedded interrogatives (WH-question or polar questions) on the one hand and embedded declaratives on the other hand. As I have shown for Hinuq and Tsez, the distinction can involve not only the use of WH-words in the complement clause (in case of embedded WH-questions), but also the use of different complementation types. In Ingush, as in many other Nakh-Daghestanian languages, embedded interrogatives do not differ in form from interrogatives in main clauses, that is, they either contain a WH-word and a finite verb, see (66b) above, or a finite verb in the interrogative form. In contrast, if the complement is a declarative (e.g. 'know that'), than the complement verb takes the subjunctive suffix; see (66c) above. Another language that formally distinguishes between embedded interrogatives and embedded declaratives is Icari Dargwa (Sumbatova & Mutalov 2003: 178–180).

Another semantic distinction, between an indicative and a modal meaning, is formally expressed in Tsova-Tush through the means of the subjunctive. Adding the subjunctive suffix *-lo* to the tense suffix in example (71), which contains a complementizer and a finite verb in the complement clause, produces the modal meaning 'that the child should stay home'.

Tsova-Tush

(71) *gadac'q'vet'aD-in-atx* [*me* *bader* *čuħ* *Dis-U*]
decide-AOR-1PL COMP child home stay-PRS
'We decided that the child is staying at home.'
(Holisky & Gagua 1994: 202)

Finally, Bagvalal expresses a further semantic distinction that is not found in the other languages of the family. The difference between genitive dependents of the masdar and preserving the valency frame is not only syntactically conditioned by the valency frame of the embedded verb; see examples (67a) and (67b) above. It also reflects a semantic distinction between a referential reading (usual valency frame) and a non-referential reading (genitive dependent). This means that if the situation denoted by the complement has a concrete spatio-temporal anchoring referring to a specific situation then all arguments of the masdar keep their normal case marking; see (72a). If, on the contrary, the complement clause does not refer to a specific situation, but rather to a situational type, then a genitive dependent of the masdar is preferable; see (72b).

Bagvalal

(72) a. *ʕali-la* *q'očan-č'i* [*Rasul-i-r* *e-w* *w-aX:a* *w-it'e:-r*]
Ali-DAT want-NEG Rasul-OBL-ERG LOG-I I-away I-send-MSD
'Ali did not like how Rasul sent him away.'
(Kalinina 2001: 526)

b. *ʕali-la* *q'oč-in-o:-b* [*ima-š:u-b* *keč'* *b-ihi-r*]
Ali-DAT want-IPFV-PTCP-III father-OBL-GEN song III-take-MSD
'Ali likes father's singing.'
(Kalinina 2001: 525)

12.4 Combinations of complementation types

Sometimes the combination of different complementation types leads to semantic differences. One example from Tsova-Tush was already given in Section 12.3, where the combination of the complementizer that can be described as semantically neutral in (71) with the subjunctive mood triggered a modal meaning.

But more commonly the combination of different strategies does not change the meaning or the mood of the complement clause. This is the case in Lezgian where the participial and the quotative particle can be used simultaneously without evoking any apparent semantic differences (Haspelmath 1993: 374). Similarly, in Icari Dargwa the enclitic *=a:l(i)* marks focus and tag questions in finite clauses and questions in complement clauses. It is encliticized to the focus of the

sentence which is commonly the predicate or a WH-word. It can be combined with the quotative particle *-ik'^{w}il* as in (73a), but neither *=a:l(i)* nor *-ik'^{w}il* are obligatory, as can be seen from (73b) and (73c). In (73b) the enclitic *=n(i)* is employed that only occurs in WH-questions. The two interrogative enclitics *=a:l(i)* and *=n(i)* cannot co-occur in one and the same clause.

Icari Dargwa

(73) a. *du-l* [*rabdan-ni* ***ce:l*** *b-a^{ʕ}rq-ib-ci* *w-**ik'^{w}il***] *x:arba^{ʕ}R-ib-da*
1SG-ERG Rabdan-ERG **what.IQ** N-do.PFV-PRET-ATR M-**QUOT** ask.N.PFV-PRET-1
'I asked what Rabdan was doing.'
(Sumbatova & Mutalov 2003: 180)

b. *du-l* [*rabdan-ni* *ce=**n*** *b-a^{ʕ}rq-ib-ci* *w-**ik'^{w}il***] *x:arba^{ʕ}R-ib-da*
1SG-ERG Rabdan-ERG what=**Q** N-do.PFV-PRET-ATR M-**QUOT** ask.N.PFV-PRET-1
'I asked what Rabdan was doing.'
(Sumbatova & Mutalov 2003: 179)

c. *du-l* [*ča* *sa:R-b=**a:l*** *x:arba^{ʕ}R-ib-da*]
1SG-ERG who come.M.PFV-PRET=**IQ** ask.N.PFV-PRET-1
'I asked who had come.'
(Sumbatova & Mutalov 2003: 179)

Finally, in Tsakhur, the modal complementizer *naXa* 'if, whether' can be combined with the neutral quotative particle *wɨ* (or the masdar), retaining the semantics of *naXa*:

Tsakhur

(74) *Rasul-u-k'le* *ac'a* *ɨxa* *deš* [*jed^{j}*
Rasul-OBL-AFFC know(IV) be(IV).PFV AUX.NEG mother(II)
*qa<r>ɨ=**naXa=wɨ***]
come.PFV<II>=**if=QUOT**
'Rasul did not know whether mother came.'
(Ljutikova & Bonč-Osmolovskaja 1999: 493)

13 Summary

In this paper I have presented a descriptive analysis of complementation in the Nakh-Daghestanian language Hinuq focusing on the semantics of complementizers and complement clauses. I have enriched the Hinuq data by comparing it to other Nakh-Daghestanian languages.

Hinuq has seven complementation strategies, but the main focus of this paper has been on the major Hinuq complementizers, namely the infinitive/pur-

posive converb, the zero strategy, the abstract suffix *-łi* and the quotative enclitic *=ǯen*. Almost all Hinuq complementation strategies make use of complementizers, which all occur as suffixes or enclitics. The quotative enclitic comes close to what can be called a canonical complementizer because it is the only one that occurs with finite complements. Neither from the point of view of their formal properties nor with regard to their semantics can the complementizers be seen as a homogenous class. Most but not all complementizers are simple non-finite verb suffixes (infinitive, participles, a masdar, and a converb). But the two most important complementizers (the abstract suffix *-łi* and the quotative enclitic *=ǯen*) do not belong to the verbal paradigm. There is usually no one-to-one relationship between complementizers and complementation strategies. Of the four semantic types available (potential type, fact type, activity type and direct speech or speech act type), most complementizers occur with two or even three of them. Some of the complementizers are predominantly epistemic and thus mostly used as modal complementizers of the fact type (e.g. *-łi*). Others are of the activity type and are not epistemic in nature (e.g. case-marked past participle or masdar). Epistemic complementizers can co-occur with other markers of epistemic modality. Only the epistemic complementizers *-łi* and *=ǯen* can be left out with some CTPs. Reported speech construction instantiate a further type of complementation, the direct speech type.

Hinuq complementizers are quite typical for the Nakh-Daghestanian languages, which generally do not have many complementizer particles of *that* or *whether* type and only very occasionally use special non-indicative verb forms such as subjunctives.

References

Abdulaev, Arsen K. & Isa K. Abdullaev. 2010. *Cezyas folklor* [Tsez folklore]. Makhachkala: Lotos.

Aikhenvald, Alexandra. 2011. Speech reports: A cross-linguistic perspective. In Alexandra Aikhenvald & Robert M.W. Dixon (eds.), *Language at large: Essays on syntax and semantics,* vol. 2, 290–326. Leiden: Brill.

Alekseev, Mixail E. & Boris M. Ataev. 1998. *Avarskij jazyk* [The Avar language]. Moscow: Academia.

Authier, Gilles. 2009. *Grammaire Kryz: Langue Caucasique d'Azerbaidjan dialecte d'Alik*. Leuven & Paris: Peeters.

Boye, Kaspar & Peter Harder. 2007. Complement-taking predicates. *Studies in Language* 31. 133–153.

Boye, Kasper & Eva van Lier. 2009. The cross-linguistic coding of the contrast between 'state of affairs' and 'proposition'. Paper presented at ALT8 conference, Berkeley, 23–26 July.

Daniel, Michael. 2013. Issues in Khinalug syntax: Building on corpus evidence. *Working Papers of the National Research University Higher School of Economics* WP BRP 27/HUM/2013. http://www.hse.ru/data/2013/05/08/1299213283/27HUM2013.pdf (accessed 7 October 2015))

Daniel, Michael & Yuri Lander. 2010. A girl of word, meat of a ram, and a life of longing: On peculiar cases of relativization in East Caucasian Languages. Paper presented at the Syntax of the World's Languages 4 conference, Lyon 23–26, September.

Dik, Simon C. 1997. *The theory of Functional Grammar,* vol. 2: *Complex and derived constructions*. Edited by Kees Hengeveld. Berlin: Mouton de Gruyter.

Dixon, Robert M. W. 2006. Complement clauses and complementation strategies in typological perspective. In Robert M. W. Dixon & Alexandra Aikhenvald (eds.), *Complementation*, 1–48. Oxford: Oxford University Press.

Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 366–431. Oxford: Oxford University Press.

Forker, Diana. 2013. *A grammar of Hinuq*. Berlin: Mouton de Gruyter.

Frajzyngier, Zygmunt. 1995. A functional theory of complementizers. In Joan L. Bybee & Suzanne Fleischman (eds.), *Modality in grammar and discourse*, 473–502. Amsterdam: Benjamins.

Frajzyngier, Zygmunt & Robert Jasperson. 1991. *That* clauses and other complements. *Lingua* 83. 133–153.

Haspelmath, Martin. 1989. From purposive to infinitive: A universal path of grammaticization. *Folia Linguistica Historica* 10. 287–310.

Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin: Mouton de Gruyter.

Haspelmath, Martin. 1996. Complement clauses. In Aleksandr E. Kibrik (ed.), *Godoberi*, 175–197. Munich: Lincom.

Haspelmath, Martin. 1999. Long distance agreement in Godoberi (Daghestanian) complement clauses. *Folia Linguistica* 33. 131–151.

Hengeveld, Kees. 1989. Layers and operators in Functional Grammar. *Journal of Linguistics* 25. 127–157.

Holisky, Dee A. & Rusudan Gagua. 1994. Tsova-Tush (Batsbi). In Rieks Smeets (ed.), *The indigenous languages of the Caucasus*, vol. 4: *The North East Caucasian languages II*, 147–212. Delmar, NY: Caravan Books.

Jendraschek, Gerd. 2009. Origin and development of the Iatmul focus construction: Subordination, desubordination, resubordination. *Folia Linguistica* 43. 345–390.

Kalinina, Elena. 2001. Aktantnye predloženija [Actant clauses]. In Aleksandr E. Kibrik (ed.), *Bagvalinskij jazyk* [The Bagvalal language], 512–553. Moscow: Nasledie.

Kazenin, Konstantin & Jakov G. Testelec. 1999. Struktura sostovljajuščix [Constituent structure]. In Aleksandr E. Kibrik (ed.), *Èlementy caxurskogo jazyka v tipologičeskom osveščenii* [Elements of Tsakhur from a typological perspective], 314–346. Moscow: Nasledie.

Kehayov, Petar. This volume. Complementation marker semantics in Finnic (Estonian, Finnish, Karelian).

Khalilova, Zaira. 2009. *A grammar of Khwarshi*. Utrecht: LOT.

Kibrik, Aleksandr E. 1994. Archi. In Rieks Smeets (ed.), *The indigenous languages of the Caucasus*, vol. 4: *The North East Caucasian languages II*, 297–365. Delmar & New York: Caravan Books.

Kibrik, Aleksandr E. 2003. *Konstanty i peremennye jazyka* [Constants and variables of language]. Saint Petersburg: Aletheia.

Klumpp, Gerson. This volume. Semantic functions of complementizers in Permic languages.
Ljutikova, Ekaterina & Anastasija Bonč-Osmolovskaja. 1999. Aktantnye predloženija [Actant clauses]. In Aleksandr E. Kibrik (ed.), *Èlementy caxurskogo jazyka v tipologičeskom osveščenii* [Elements of Tsakhur from a typological perspective], 481–536. Moscow: Nasledie.
Nichols, Johanna. 1994. Chechen. In Rieks Smeets (ed.), *The indigenous languages of the Caucasus*, vol. 4: *The North East Caucasian languages II*, 1–77. Delmar & New York: Caravan Books.
Nichols, Johanna. 2011. *Ingush Grammar*. Berkeley & Los Angeles: University of California Press.
Polinsky, Maria. 2003. Non-canonical agreement is canonical. *Transactions of the Philological Society* 101. 279–312.
Polinsky, Maria & Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19. 583–646.
Sumbatova, Nina R. & Rasul O. Mutalov. 2003. *A grammar of Icari Dargwa*. München: Lincom.
Thompson, Sandra A. 2002. "Object complements" and conversation: Towards a realistic account. *Studies in Language* 26. 125–164.
van den Berg, Helma. 2005. The East Caucasian language family. *Lingua* 115. 147–190.
van Lier, Eva & Kasper Boye. 2010. Epistemic complementizers: A cross-linguistic survey. Talk at the University of Leipzig, 14 January 2010.

Peter Harder

Complementizers and the complementizing function – structure and pragmatics in a panchronic perspective

1 Introduction

There are several related aims with this paper. First, I try to show how the basic nature of complementation can be illuminated by seeing the diachronic rise of complementation as an instance of grammaticalization understood as an *ancillarization* process: grammatical elements have a secondary and ancillary semantic role, rather than being carriers of primary meanings (cf. Boye & Harder 2012). What makes complementation a key instance of the significance of the rise of new grammatical elements is the fact that their secondary, linking role is closely tied to the design feature of human languages that is associated with recursion; recursion comes about when clauses become structurally instead of only pragmatically integrated.

The account provided also reflects a second point of principle: to show the importance of a *content-syntactic* approach to clause structure and complementation, i.e. an approach (cf. Engberg-Pedersen et al. 1996) where hierarchical relations between conventional meanings are structurally central. Among such relations, the top-down relation of *function-assignment* is especially crucial for understanding the nature of complementizers.

Based on these two chief theoretical concepts, the more descriptive aims of the article are to account for the semantic niche of complementizers in the light of the synchronic functional pressures that are at work in complementation, and finally to illustrate this panchronic approach with reference to some contested issues in relation to Danish complementizers.

2 The theoretical significance of complementation

It is generally accepted that human language has the unique design feature of structural complexity. In accounting for this feature, generative grammar has attributed special significance to recursion as perhaps the genetically unique component of the human language ability to make infinite use of finite resources (cf. Hauser, Chomsky & Fitch 2002). While the question of the precise status

of recursion is controversial (cf. van der Hulst 2011), the mechanism whereby one clause can become subordinate to another remains a particularly striking example of the creation of more complex structures out of simpler elements.

In functional and cognitive linguistics the question of complementation has given rise to a major theoretical issue, arising partly as a result of scepticism with respect to the emphasis on purely structural complexity in generative grammar. The question is whether formally subordinate clauses are really subordinate – or, instead, they contribute the main semantic content and the so-called higher clause is better understood as an appendix with particular semantic functions (cf. Thompson 2002; Newmeyer 2010). However, the present article assumes (cf. Boye & Harder 2007) that there is a key type of case where structural and semantic subordination coincide, so that one clause is assigned a role as one part of a more complex clause, which then constitutes the larger, superordinate semantic unit. The analysis refers to such cases, without denying the existence of other types.

In terms of this picture, complementizers must be understood as belonging at the top end of the semantic hierarchy, or put differently, of the content-syntactic structure of the clause. Structurally, this is in accordance with their role as functional heads in generative clause structure, but in addition to this, the theory presented involves an assumption that they have conventional meanings that constitute the foundation of this structural position. Hence, the theory is in conflict with the familiar assumption that complementizers are semantically empty markers of structure. The difference can be expressed by saying that complementizers could not mark structure, as indeed they do, if they had no semantic content – and that their content is responsible for the structure they mark, without being reducible to it.

In short, complementizers are central to the issues of structural complexity and the relation between structure and meaning. Thus they are suitable points of departure for addressing this issue, including the relation between item meanings and structural relations, hierarchical as well as constructional, and also the relation between pragmatic and conventional meaning.

3 Complementation as a product of grammaticalization

The synchronic issue of structural complexity must be understood as one half of the full panchronic picture that also includes the diachronic rise of structural properties. Among these, the most basic property is arguably that of being a

grammatical rather than lexical element in the language. There is no consensus about what that precisely entails, but based on Boye and Harder (2012) an argument can be presented for why an understanding of the nature of grammatical status can throw light on the nature of complementation.

In this picture, a grammatical element is one which plays second violin to another expression, adding ancillary content that enhances the functional potential. This approach thus demonstrates how grammatical elements can be less prominent than lexical expressions and still add something essential – by enhancing the powers of the whole complex expression of which they form part.

Phylogenetically speaking, full-blown grammaticalization is a two-step process, cf. Boye & Harder (2012), of which only the first step is obligatory: The first step is when a pragmatic relation between expressions becomes conventionalized as a structural relation: *Mummy! There!* > *Mummy there!* This is a process of *constructionalization*, in that the constituent items give rise to something that is not reducible to the elements in themselves, of which the elements form structural parts. Grammatical meaning in such cases is in the nature of constructional meaning, in this stipulated case taking the form of the proto-predicational or *pivot-operator* relation (cf. Braine 1963).

The most familiar form of grammaticalization in this theory constitutes a possible second step (delexicalization) – when a lexical expression gets recruited to serve a secondary function in relation to another expression: *he is **going** (in order) **to** help* > *he's **gonna** help*. It follows from this theory that grammaticalization always creates a new *construction*, i.e. a new and more complex relation between input items (with both expression-side and content-side aspects), while it may – but need not – create a new grammatical item that enters into the new semantic complex.

In the case of complementation, this account of the nature of grammaticalization can be illustrated with two possible pathways:

(1) The python pathway (Hidatsa)

(2) The demonstrative pathway (English, Faroese etc.)

The python pathway creates a construction where a sentence is swallowed up by its neighbour, without involving the rise of new items.

In Hidatsa, all sentences end with a mood operator (Matthews 1965: 55; 98). 'Say'/'think/'know' verbs may swallow up a whole independent clause (or even two adjacent clauses) as objects. These retain their own final mood markers – but

the complex is then capped by an extra mood morpheme (quotative) with scope over the whole sentence.[1]

In this way, a complex construction arises in which one clause is structurally subordinate to another, without giving rise to a new item that signals this relationship. This clearly is an aspect of the grammar of the language, and the process whereby it came about thus qualifies as an instance of grammaticalization. This python pathway represents what above was called the first step in a full theory of grammaticalization, the rise of a new construction. What in (2) I called the demonstrative pathway exemplifies the more complex, full-blown process in which a new ancillary *item* comes into being as a key element in the new construction. This pathway, which is known from English, German, Faroese and other languages (cf. Fischer 2007), involves a transition that can be roughly illustrated as follows (cf. also Boye & Harder 2012):

(3) *He said **that**: You are nuts > He said that you are nuts*

The input consists of two independent clauses, linked by a non-conventionalized, pragmatic relationship: the demonstrative in the first clause refers to the content of the second clause. The analysis of this relationship as purely pragmatic can be illustrated by comparison with an utterance of *he said that* in which there is no second clause, so that the demonstrative refers situationally to something that has previously been mentioned by another speaker.

A scenario for the process of change would then be a gradually increasing frequency (cf. Bybee 2006) of the sequence in which a demonstrative is followed by a clause specifying the content to which the demonstrative refers. Such a sequence would involve competition between the demonstrative and the clause as filler of the object slot of *said*: the original parsing strategy in which the demonstrative fills the slot gets a potential competitor in the form of a strategy that skips the demonstrative and goes straight to the clause that provides the content of what was said.

If the demonstrative gradually loses out in this competition, two things happen:

(4) The second clause acquires a grammatical relation with the first (constructionalization)

1 In order for this to be understood as a case of grammaticalization in the strict sense, we need to make the assumption that simple clauses came diachronically first. The grammaticalization in question would then be the subsequent rise of the construction in which adjacent clauses came to be included (swallowed up) as objects in a complex sentence.

(5) The demonstrative becomes an ancillary *linker* that introduces the "real" object: *he said that you are nuts*

Viewed from a language situation in which clauses can only be linked pragmatically rather than grammatically, this is a development fraught with all sorts of interesting consequences, as discussed in the introduction. But it should be pointed out that in itself it is also simply the rise of a new construction with a special item signalling the new relationship between the component parts. Only when the same pattern is replicated in other developments can we confidently speak of the rise of complementizers as a new general category of expressions.

4 Item semantics and content syntax

As suggested above, the framework of analysis involves an understanding of clause grammar in which content syntax is the backbone of syntactic relations. The key assumption is that coded, conventional meanings have syntagmatic, including hierarchical relations, cf. Engberg-Pedersen et al. (1996, 2005)

The basic layout of content-syntactic relations in the clause is common to various forms of functional syntax (Foley & van Valin 1984, Dik 1989, etc.), with a skeleton roughly paraphrasable as (6):

(6) grounding (qualification [predicational core]).

At the top end of the hierarchy, we find *clausal illocution*, which takes the rest of the clause content inside its scope, and specifies what speech act type it serves to perform. This is a clause-structural manifestation of Searle's (1969) distinction between propositional content and illocutionary function, F(p). In a philosophical context, a large number of different illocutionary functions can be recognized, but only a few abstract illocutionary categories are structurally encoded, with *declarative* as the most salient case – the abstract nucleus of all speech act types that involve the speaker's commitment to the truth of the proposition.

This illocutionary element assigns a function to the content inside its scope – in this case a speech act function that the propositional content is recruited to serve. But function-assignment can be generalized (cf. Harder 1996a: 79; 150) to include other functions than independent speech act functions, applying generally to top-down semantic relationships, also at lower levels. In the European tradition, it is customary to analyse all clause constituents in terms of two levels, the functional level (e.g. subject, modifier, adverbial etc.) and the level of material (noun phrase, prepositional phrase, adjective, etc.); the same pattern of think-

ing is reflected in the slot-filler-distinction. The process whereby expressions in a clause combine to produce a coherent whole requires that the proper function is assigned to each constituent, and this is partly achieved by having markers that assign functions to constituents; for instance, a dative affix may assign the function of indirect object to a noun phrase.

The mechanism of function-assignment is crucial to understanding the content-syntactic niche in which complementation belongs. When a clause is inserted into a complementizing frame, the key result is that the clause is subjected to a *reassignment of function*. A clause that comes inside the scope of complementation no longer serves as an independent declarative speech act, but gets recruited to serve a dependent function within a larger whole. Consider the difference between (7) and (8):

(7) *he is dead*.

(8) *that he is dead*.

Standing on its own as in (7) *he is dead* commits the speaker to the truth of the proposition; prefixed by *that*, as in (8), however, it belongs in a dependent function as part of a more complex statement, cf. the alternative complement-taking frames *I believe*... and *I hope*..., neither of which commits the speaker to its truth. The same re-assignment occurs in the case of adverbial subordinators such as *because* and *if* (cf. Harder 1996b): prefixing a clause with *if*, as in *if he is dead* converts the proposition to condition on the main clause speech act coming into force, as in *if he is dead, tell me at once!*

As an argument for why this parallel between the assignment of illocutionary function and the assignment of intra-clausal functions is not linguistically fanciful, it can be pointed out that in some languages (like West Greenlandic, cf. Fortescue 1984, and Hidatsa, cf. Matthews 1965), certain subordinators occupy the same paradigmatic slot as independent illocutionary operators. *Contemporative* function (~ 'while') may thus constitute a structural alternative to declarative function. In a very different functional domain, the same general type of reassignment can be illustrated with a bank manager whose company is bought by a larger bank; he is no longer the head of an independent business, but may continue as branch manager, lower in the structural hierarchy.

In none of these cases is this reassignment of function an empty and meaningless operation. For the same reason, the fact that there is a difference (in Danish) between word order in main and subordinate clauses is not semantically empty (cf. Hansen & Heltoft 2011). For a key class of cases, it marks roughly whether the clause content should be assigned an independent declarative function or it is a subordinate semantic foot soldier.

If we look only at item meaning, however, there is little similarity between illocutionary-type meanings (declarative, interrogative etc) and the meanings of complementizers and subordinators (*that, while, if*). The parallel is strictly in content-syntactic role – and only if viewed top-down, i.e. in relation to the conventionally assigned function of the proposition that is being considered. Viewed bottom-up, the complementizer is at a lower level in the hierarchy than the independent illocution (branch manager vs. CEO), hence not in the same semantic environment.

This complex relationship between complementizers and subordinators on the one hand and the structural-semantic hierarchy on the other gives rise to a number of semantic pressures that influence what item meanings are possible for the class of items that fill this slot. The meanings of complementizers and subordinators must therefore be understood in relation to the semantic pressures that are built into their content-syntactic position: Upwards, they have to fit into the slot defined by the matrix clause – and downwards, they have to assign a function to the clause in their scope.

On this basis, I would like to offer my proposal for what we should take the complementizing function to be. Simply put, it is to assign to a clause the function of serving as a constituent in a(nother) clause. In the narrow sense, the role in question is as a core argument of the clausal verb. In the sense that includes subordinators, it includes a wider range of clausal functions. The function-assignment can work either purely by inserting the clause in a constructional slot, or it can involve a dedicated function-assigning grammatical item (a complementizer).

5 Complementation, subordination and pragmatic linkage

Above, we have seen how various structural (more precisely, structural-semantic) factors play a role for understanding complementizer semantics. We now turn to those pragmatic mechanisms that interact with the structural-semantic factors.

The basic nature of the pragmatic dimension can be understood in relation to the Gricean principles. Just as an utterance is understood in a way that allows it to be seen as co-operating with what is going on in the pragmatic context, so the meaning of a conventional expression must be understood in a way that enables it to co-operate with its syntagmatic context (cf. *red hair* vs. *red scare*). The phenomenon of coercion is one manifestation of this general principle (cf. Harder 2010: 245–50), applying to the form of collaboration that consists in a

higher-scope meaning imposing a particular reading upon the meaning of an expression in its scope (in *a beer* the indefinite article thus imposes individuation upon the substance denoted by the noun *beer*, cf. Michaelis 2003).

Since the top-down dimension of complementation has been highlighted above, this makes coercion relevant to the understanding of complementizer semantics. Coercion works at two levels in relation to complementation: (1) from the matrix clause slot, operating on the combination of complementizer and subclause; (2) from the complementizer, operating on the subclause. At level (1), coercion constricts the semantic freedom of function assignment by the complementizer – most severely for complementizers in the strict sense (cf. Dixon 2008): those that assign the function of filling argument slots, prototypically the object slot. Coercion to fit into the object slot has two semantic effects:

– reification (in order to fill a nominal position)
– adjustment to selectional restrictions (as in *Joe greeted the pillar-box*, which forces a reading in which animacy is attributed to the pillar-box, or as in factivity: *I regretted that I did it* forces the subclause into the role of conveying a fact).

Coercion from the matrix clause may make the complementizer *that* redundant if the object role of the subclause is signalled in other ways – i.e. in English and Danish if the clause comes after the verb, rather than in the case of initial *that*-clauses. Since *that* is optional in such cases, it is logical to assume that reification is both what occurs when a clause is coerced into object function (by occurring after a complement-taking verb) and also the conventional meaning of the complementizer *that*. You can achieve the same semantic effect in either way, and you can also both use *and* and have the clause in post-verbal position.

However, reification is not all there is to say. There is some residual functional freedom in the relation between subclause and matrix clause, including a cline from relative independence to complete nominalization. Among the more striking examples of this is the fact that Danish allows subclauses to have full declarative force, as signalled by main clause declarative word order, cf. Hansen & Heltoft (2011: 1685):

(9) *Jeg sagde* [***at*** *det var ikke ham*]
I said **that** it was not him
'I said that it was not him', with the subclause marked as having full declarative force

Subordinate word order would have the negation before the verb:

(10) *Jeg sagde* [***at*** *det ikke var ham*.]
I said **that** it not was him
'I said that it was not him', with the subclause marked as a dependent constituent

More generally, semantic freedom also depends on the specificity of the meaning (and hence of the selectional restrictions) of the matrix verb: a more general verb like *say* gives more leeway than *ask* and *wonder*.

The relative freedom of complementizer-semantic options in subclauses after 'say'-type verbs can be illustrated with a development in Korean. The basic nature of the options available can be understood as a case of main clause choices being "downloaded" into the subclause, instead of the subclause being understood merely as a string of quoted verbs, cf. Rhee:

> [the] quotative marker underwent a fusion with the sentential ending of the embedded direct speech, which varied depending on the sentence type, i.e., *-ta* for declarative (...) *-nya* for interrogative, *-la* for imperative, and *-ca* for hortative, (...) This process of morphological fusion brought forth a complete paradigm of complementizers with four members depending on the sentence type of the reported utterance. (Rhee 2009: 203).

Such a development does not clarify the precise synchronic distribution of semantic labour between matrix clause coercion and complementizer and subclause semantics at different stages of development. At the same time, it illustrates the non-neat separation of factors at work in the syntagmatic space defined by the main clause at the higher level and the subclause inside the scope of the complementizer: How much work there is to do for the complementizer is variable, although always somewhat restricted.

If we turn to adverbial subordinators, the same variability is at work. Some subordinators take full declarative illocutions inside their scope (e.g., *because*, as in *he left because he was ill*). Others assign inherently subordinate clause functions (e.g., *if* as a condition of the main illocution coming into force, cf. Harder 1989). A particularly interesting case of the relation between subordination and illocutionary function-assignment occurs when subordinate clauses break out of their subordination and upgrade (by pragmatic implicature) to minor illocutionary types. Thus clauses introduced by *if only*...(also known as insubordination) have come to signal 'wish' when they stand without what used to be a following main clause.

Finally, while the cases discussed are understood as instances of the situation where structural subordination is basically correlated with semantic subordination, I would like to make explicit once more that this is not always the case. Matrix clause complement-taking predicates may grammaticalize, cf. Boye & Harder (2007), leaving them as secondary accompaniments to what originated

as a complement clause in ways described in detail by Thompson (2002) – clear cases being formulae like *I bet* or *I guess*.

6 Complementation in Danish – and the status of *hvis*

One of the types of meaning that can squeeze into the semantic space left open for complementizer items is epistemic meaning (cf. Boye 2008; Boye, van Lier & Brink 2015). This is natural because epistemic meaning is close to the illocutionary top of the content-syntactic hierarchy. It is also a type of grounding meaning in that it signals not objective properties but the epistemic status assigned to the utterance by the speaker.

As an example, Jacaltec (Craig 1977, as quoted by Boye 2008: 26) offers a choice between two complementizers, one for certainty (*chubil*) and one for uncertainty (*tato*), in *say*-complements. By marking a statement with *chubil*, the speaker reveals his own epistemic stance (as speaker) to information from a different source (thus co-grounding the claim). Apart from being conventional meaning, epistemic stance may also occur as pragmatic, i.e. situational or implicated information. Japanese *-tara* and German *wenn* may be interpreted as *when* if the speaker is certain, otherwise as *if (*cf. Boye 2008).

The two types of meaning (epistemic and illocutionary) are so close in the semantic hierarchy that the same expressions often have implications on both levels. For instance, asking a yes/no question (an illocutionary function) presents the propositional content as something the speaker is epistemically uncertain about (an epistemic function). A theoretical analysis can account for this in terms of functional dependence, cf. Boye (2012: 308–315), but in concrete cases the analysis may not be unambiguously decidable. Report is a form of evidentiality, and at the same time to give a report is a type of speech act. Moreover, in the somewhat cramped semantic space of the complementizer slot, differentiation is inherently limited, so it is to be expected that there will be situations when it is not obvious which type of being is basic: There are cases when complementizers can plausibly be assigned (sub)illocutionary-type meaning, while equally plausibly this meaning may be understood as epistemic. One may venture the conjecture that epistemic stance may offer the best available locus of generalization about complementizer semantics (cf. Boye, van Lier & Brink 2015) because it does not take up the illocutionary turf that is typically not fully available in the complementizer slot.

However, the complex forces that are at work in the territory means that the issue is difficult to clinch. As a case in point, consider the analysis in Boye (2008) of the Danish complementizers *at* (= 'that') and *om* (= 'whether') as coding neutral vs. uncertain epistemic stance. This contrasts with an earlier widespread analysis of *om* as indicating interrogative (sub-illocutionary) status. This account also involves a neat division of labour between *hvis* (='if') as adverbial indicator of uncertainty and *om* ('whether') as nominal indicator of uncertainty.

Nevertheless, there may be minor Procrustean consequences of this analysis. The possible collateral damage involves the arguably sub-illocutionary aspects of the meaning of *om* and *hvis*.

To begin with, *om* differs from Jacaltec *tato* in that it cannot be used to indicate that a reported source is not fully reliable, as in cases like 'He said *tato* the president is coming'. The putative corresponding Danish version with *om* in the position of *tato* would be *han sagde om præsidenten kom* [3SG say.PST COMP president-DET come.PST], which would mean 'he said whether the president is coming' (i.e. yes or no). As an indicator of uncertainty, it can only be used to fill slots (object after *ask, doubt*, etc) that naturally suggest that the truth value is undecided (which is not quite the same as expressing speaker uncertainty).

This could be expressed by saying that *om* assigns the proposition the (sub-illocutionary) status of an *issue* rather than a potential fact. Issues are things you need to decide on in order to take action, not dodgy potential truths. Therefore presenting something as an issue does not have to signal speaker uncertainty about truth. The teacher who asks his students *how does one determine whether the earth moves round the sun or not?* does not thereby signal uncertainty, only that they should address this issue. Yet in most cases there is a very thin line between presenting something as an issue and expressing uncertainty about it, so the case illustrates both the epistemic potential and the possibility of differentiation.

A similar argument that sub-illocutionary function rather than epistemic status may be the shared *grundbedeutung* can be offered for the other, adverbial marker of epistemic uncertainty. Like its English counterpart *if*, the Danish subordinator *hvis* assigns the sub-illocutionary status of being a condition on the coming into force of the matrix clause speech act (*if it rains, matches are played indoors)*: *it rains* is not a proposition whose truth you wish to assert, but one that if true licenses a speech act that you are not otherwise in a position to perform; this speech act does not have to be a statement (compare. *if it rains, go indoors!*) cf. Harder (1996a: 446–447). Since *if* also in this function leaves the truth value

open, this may naturally signal epistemic uncertainty (as is clearly the case for the nominalised sense: *there are too many ifs!)*.[2]

The conditionality reading of *if* that is standardly associated with adverbial *if*-clauses is in some cases compatible also with its occurrence in the complementizer slot (cf. also Mortensen 2010):

(11) *Jeg hader* [***hvis*** *folk har et problem med min hårfarve (blondine!!)*]
I hate **if** people have a problem with my haircolour (blonde)
'I hate if people have a problem with my hair colour (blonde!!)'
(http://www.synes-godt-om.dk/8252)

In this position, *hvis* contrasts with the complementizer *at*, which would give a factive reading. Moreover, in such contexts, *hvis* cannot be replaced by *om* in Danish. But there are also contexts where all three items (*at, om, hvis*) are possible:

(12) *Det er vigtigt, at du fortæller,* [***hvis*** *du har smerte-r*]
It is important that you tell **if** you have pain-PL
'it is important that you tell if (=in case) you are in pain' (notice in the coronary ward).
(http://www.hvidovrehospital.dk/NR/rdonlyres/0535B5B2-A52D-45B6-8A00-E1FBA3057F62/0/VelkommentilHjerteafdelingen115.pdf)

2 Boye (p.c.) suggests that the relation of conditionality between *if*-clauses and main clauses can most adequately be understood as the result of a combination between uncertainty as signalled by *if* and the circumstantial relation that *if*-clauses share with other adverbial clauses (such as *when*-clauses). Circumstantial clauses all have the role as truth-conditions or validity conditions on the speech act conveyed by the main clause, cf. *if/when in doubt, leave it out*): hence, no extra sub-illocutionary meaning of conditionality is necessary to account for this conditional relationship. I recognize the force of this argument and at present I see no knock-down argument for deciding the issue of what conventional meaning should be assigned to *if*: whether uncertainty is a by-product of conditionality or conditionality is the result of a combination of uncertainty plus circumstantial relation. But this argument about conditional *if* and its Danish counterpart *hvis* can be kept separate from the more general issue of sub-illocutionary meaning as part of the understanding of the semantics of complementizers. I argue that sub-illocutionary meaning is part of the picture, but whether in specific cases it should be understood as implied or conventional is not crucial. Clearly a conditional relation is sometimes purely the result of inference (as in the slogan for American tort lawyers *we don't win – you don't pay)* and sometimes signalled as a by-product of expressions whose conventional meaning do not include conditionality (*do that, and I'll kill you*). Whether a sub-illocutionary reading is conventionalized or inferred needs to be investigated separately for each case.

(13) *Det er utrolig vigtigt at du fortæller* [***om*** *du synes det er godt*] *eller dårligt*
It is incredibly important that you tell **whether** you think it is good or bad
'it is incredibly important that you tell whether you think it is good or bad'
(http://danskfestmusik.dk)

(14) *Du fortæller,* [***at*** *du skændes med dine forældre*]
You tell **that** you quarrel with your parents
'you say/tell that you quarrel with your parents'
(www.dr.dk/stopselvmord)

The putative *hvis*-complement clauses can be analysed as abbreviated versions of clauses that include a pronominal object, such as *it is important that you tell [it] if* ... However, in a panchronic perspective this can be true without necessarily ruling out the simultaneous relative validity of the *hvis*-complement analysis. The process that elides the pronoun is potentially analogous to an aspect of the postulated historical development that was at work when the demonstrative *that* changed into the complementizer *that*: we assumed, as discussed in relation to (2), that the pronoun was overruled in order to establish a direct relation between the verb and the following clause. There is a difference in that the pronoun does not turn into a complementizer but is rather left out – but this can be accounted for by reference to the fact that the following clause already has a subordinating element (*hvis*), so there is no slot for the 'it' to fill, once a direct relation has been established – in (10), between *fortælle* ('tell') and *hvis du har smerte-r* [COMP 2G have.PRS pain-PL] ('in case you are in pain') – it is simply squeezed out because the following *hvis*-clause is being sucked into the complement position. To the extent this is the case, *hvis* is being conventionalized as a complementizer in the narrow sense rather than a subordinator.

This case also illustrates the way in which the semantic potential of *hvis* is redeployed in a way that reflects the combined structural and pragmatic pressure of the intensified coercion that is at work in the pure complementizer position. Here, the object function leaves less room for the conditional subillocution, which is standardly associated with the circumstantial relation that is characteristic of adverbial clauses (cf. note 2 above): the complementizing verb needs an object entity to operate on, so the reifying pressure submerges the conditional meaning. As a result, the epistemic dimension stressed by Boye (2008) comes into its own in this position; the difference between *I hate that*... and *I hate if*..., cf. (9), can be captured exhaustively by the contrast between factual and uncertain epistemic status.

7 Conclusions

Complementizers are structural elements, and as such they occupy one of the most theoretically central roles in the structure of human languages. This account has tried to show that at the same time they can only be adequately understood in a panchronic perspective that integrates their diachronic emergence from pragmatic relations between clauses, their synchronic position in the content-syntactic hierarchy of the clause, and their dependence on Gricean mechanisms of semantic adjustment, including coercion.

One diachronic aspect of this account is its application of the ancillarization theory of grammaticalization to the understanding of complementizers, which as a side effect provides an account of the relation between complementation with and without a special class of complementizer items.

The overall logic of this theory was applied to some contested issues in the theory of Danish complementizers, illustrating how the same pragmatic forces that are at work in the emergence of demonstrative-style complementizers in the first place may also be at work in relocating an adverbial subordinator so that it can also take on a purely complementizing function.

Acknowledgment

I would like to thank an anonymous reviewer and Kasper Boye for valuable comments and suggestions. Needless to say, remaining errors are entirely my own.

References

Boye, Kasper. 2008. Om om [About om]. *Ny forskning i grammatik* 15. 7–29.

Boye, Kasper. 2010a. Raising verbs and auxiliaries in a functional theory of grammatical status. In Kasper Boye & Elisabeth Engberg-Pedersen (eds.), *Language Usage and Language Structure*, 73–104. Berlin: De Gruyter Mouton.

Boye, K. 2010b. Reference and clausal perception-verb complements. *Linguistics* 48(2). 391–430.

Boye, Kasper. 2012. *Epistemic Meaning: A Crosslinguistic and Functional-Cognitive Study*. (Empirical Approaches to Language Typology 43). Berlin & New York: Walter de Gruyter.

Boye, Kasper & Peter Harder. 2007. Complement-taking predicates: usage and linguistic structure. *Studies in Language* 31(3). 569–606.

Boye, Kasper & Peter Harder. 2009. Evidentiality: Linguistic categories and grammaticalization. *Functions of Language* 16(1). 9–43.

Boye, Kasper & Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88.1. 1–44.
Boye, Kasper, Eva van Lier & Eva T. Brink. 2015. Epistemic complementizers: A crosslinguistic survey. *Language Sciences* 51.5. 1–17.
Braine, Martin D.S. 1963. The Ontogeny of English Phrase Structure: The First Phase. *Language* 39. 1–14.
Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82. 529–551.
Craig, Colette G. 1977. *The structure of Jacaltec*. Austin: University of Texas Press.
Dik, Simon C. 1989. *The Theory of Functional Grammar. Vol 1: The Structure of the Clause*. Dordrecht: Foris.
Dixon, Robert M. W. 2008. Complement clauses and complementation strategies in Typological Perspective. In Robert M. W. Dixon & Aleksandra Y. Aikhenvald (eds.), *Complementation. A Cross-Linguistic Typology*. Oxford: Oxford University Press.
Engberg-Pedersen, Elisabeth, Michael Fortescue, Peter Harder, Lars Heltoft & Lisbeth Falster Jakobsen (eds.). 1996. *Content, Expression and Structure: Studies in Danish Functional Grammar.* Amsterdam: John Benjamins.
Engberg-Pedersen, Elisabeth, Michael Fortescue, Peter Harder, Lars Heltoft, Michael Herslund & Lisbeth Falster Jakobsen. 2005. *Dansk funktionel lingvistik*. Copenhagen and Roskilde: Copenhagen University and Roskilde University.
Fischer, Olga. 2007. *Morphosyntactic change: Functional and formal perspectives*. Oxford: Oxford University Press.
Foley, William A. & Robert D. Van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge Univer¬sity Press.
Fortescue, Michael. 1984. *West Greenlandic*. London: Croom Helm.
Frajzyngier, Zygmunt & Robert Jasperson. 1991. *That*-clauses and other complements. *Lingua* 83. 133–153.
Hansen, Erik & Lars Heltoft. 2011. *Grammatik over det danske sprog* [Grammar of the Danish language]. Odense: Syddansk Universitetsforlag.
Harder, Peter. 1989. The instructional semantics of conditionals. *Working Papers in Functional Grammar* 30, 1–42. Amsterdam: University of Amsterdam.
Harder, Peter. 1996a. *Functional Semantics*. Berlin & New York: Mouton de Gruyter.
Harder, Peter. 1996b. Subordinators in a semantic clause structure. In Betty Devriendt, Louis Goossens & Johan van der Auwera (eds.), *Complex structures. A functionalist perspective*, 93–118. Berlin: Mouton de Gruyter.
Harder, Peter. 2010. *Meaning in Mind and Society. A Functional Contribution to the Social Turn in Cognitive Linguistics*. Berlin & New York: Mouton de Gruyter
Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The Faculty of Language: What is it, who has it, and how did it evolve? *Science* 298. 1569–1579.
Katlev, Jan. 2000. *Etymologisk ordbog. Danske ords historie* [Etymological dictionary. The history of Danish words]. Copenhagen: Politiken.
Kawasha, Boniface. 2007. Subject-agreeing complementizers and their Functions in Chokwe, Luchazi, Lunda, and Luvale. In Doris L. Payne & Jaime Peña (eds.), *Selected Proceedings of the 37th Annual Conference on African Linguistics*, 180–190. Somerville, MA: Cascadilla Proceedings Project.
Kirsner, Robert S. & Sandra A. Thompson. 1976. The role of pragmatic inference in semantics: A study of sensory verb complements in English. *Glossa* 10. 200–238.

Mann, William C., Christian M. I. M. Matthiessen & Sandra A. Thompson. 1992. Rhetorical Structure Theory and Text Analysis. In William C. Mann & Sandra A. Thompson (eds.), *Discourse Description: Diverse linguistic analyses of a fund-raising text*, 39–78. Amsterdam: John Benjamins.

Matthews, George Hubert. 1965. *Hidatsa Syntax*. The Hague: Mouton & Co.

Mortensen, Sune. 2010. Problemer og udveje i dansk ledsætningsklassifikation [Problems and solutions in Danish subordinate clause classification]. Paper presented at the symposium for Dansk Netværk for Grammatik, 26–27 October.

Michaelis, Laura. 2003. Headless Constructions and Coercion by Construction. In Elaine J. Francis & Laura A. Michaelis (eds.), *Mismatch: Form-Function Incongruity and the Architecture of grammar.* Stanford: CSLI Publications.

Newmeyer, Frederick J. 2010. What conversational English tells us about the nature of grammar: A critique of Thompson's analysis of object complements. In Kasper Boye & Elisabeth Engberg-Pedersen (eds.), *Language Usage and Language Structure*, 3–43. Berlin: De Gruyter Mouton.

Rhee, Seongha. 2009. Through a Borrowed Mouth: Reported Speech and Subjectification in Korean. *Lacus Forum* 34. http://www.lacus.org/volumes/34/217_rhee_s.pdf (accessed 30 August 2011).

Searle, John R. 1969. *Speech Acts*. Cambridge: Cambridge University Press.

Thompson, Sandra A. 2002. 'Object complements' and conversation. *Studies in Language* 26(1). 125–164.

van der Hulst, Harry (ed.). 2011. *Recursion in Human Language*. Berlin: De Gruyter.

Verhagen, Arie. 2005. *Constructions of Intersubjectivity. Discourse, Syntax, and Cognition*. Oxford: Oxford University Press.

Petar Kehayov and Kasper Boye

Complementizer semantics in European languages: Overview and generalizations

1 Introduction

The introductory chapter of this book listed the following questions pertaining to complementizer semantics.

A. How common is it for complementizers to have semantic functions?
B. Which kinds of semantic functions may complementizers have?
C. Which kinds of semantic contrasts are found in complementizer systems?
D. How do semantic complementizer functions develop diachronically?

As mentioned in the introductory chapter, the answers to these questions may cast new light on the following topics:

a. Complementizer optionality
b. Complementizer combination
c. Subordinator functions
d. Modality and other notional categories

In this chapter we try to get closer to answers to questions A–D, and to summarize what we can learn about issues a–d, based on the preceding chapters on complementizer semantics in individual European language families. For a number of reasons the generalizations and overviews we propose are at best tentative. Firstly, even though the contributors to this book were asked to address a number of specific issues (cf. the introductory chapter), there is considerable variation in how and to what extent the issues are (and can be) dealt with. Secondly, views on how to best describe or account for a given phenomenon differ. In the present chapter we do not always follow the analyses proposed in the preceding chaptes. Occasionally, we base our generalizations and overviews on analyses that differ from those made in the chapters dealing with individual European language families or their branches. It may turn out that we were wrong in doing so. Thirdly, crosslinguistic comparison arguably presupposes a tertium comparationis in the form of a set of relevant comparative concepts (see Haspelmath 2010 on this notion), but there is no such thing as a set of comparative concepts agreed upon by everyone. There may be disagreement about which of two candidates for a comparative concept is more adequate, and there may be disagreement about which of two relatively well-established comparative concepts is more adequate

for capturing a given phenomenon. In this chapter we generally use comparative concepts that are widespread and well-known to most linguists (as in the case of "evidentiality", and "time"; cf. sections 2.2.2 and 2.2.4). For less well-known concepts, we provide definitions (as in the case of "proposition" and "state-of-affairs"; cf. Section 2.1.1).

The chapter is structured as follows. In Section 2 we discuss the kinds of semantic complementizer functions and contrasts found in indigenous European languages (thus addressing questions A–C and issue d above). In Section 3 we deal with complementizer optionality, and in Section 4 with combinations of complementizers (thus addressing respectively issues a and b). In Section 5 we give an overview of the various kinds of functions complementizer forms may have that are distinct from their complementizer uses (thus addressing issue c above), and in Section 6 we give a survey of the ways in which semantic complementizer functions may develop diachronically (thus addressing question D).

Wherever possible, we adhere to the following terminological practice: by *complementation marker* we refer to all means (including in principle word order and intonation) of identifying complements as complements; by *complementizers* we refer to conjunctions (whether words, clitics or affixes) that serve to identify complements as complements; by *canonical complementizer* we refer to the subset of such conjunctions that serve to identify balanced (in the sense of Stassen 1985) complements. Thus, canonical complementizers are a subset of complementizers, which in turn are a subset of complementation markers.

2 Semantic types

Frajzyngier (1995) claimed that all complementizers are modal (see Nordström 2010 for a recent, modified restatement). In contrast, Boye, van Lier & Brink (2015) found, in a survey of complementizers in 89 languages, that not all complementizers are modal, but rather may have only complementizing function (i.e. the function of identifying complements as complements) or have complementizing function plus a non-modal function. The material presented in this book supports Boye, van Lier & Brink. At least three of the 46 languages covered seem to have a complementizer with only complementizing function: Hungarian (*hogy*), Kildin Saami (*šte*) and Kalmyk (*giž* and variants thereof)[1]. In addition, some lan-

1 Knyazev (this volume) analyses this complementizer as indicating non-factivity, as it is found after non-factive verbs. However, it is found also after factive emotives, as well as after directive utterance verbs.

guages have complementizers with non-modal functions. Still, as will be clear from what follows, there is a strong tendency for complementizers to have functions that are modal in a more or less strict sense of this notion.

In this section we give an overview of the types of semantic functions and contrasts that are found in the complementation markers in indigenous European languages, dealing first with frequent semantic types and then with less frequent ones. We discuss semantic contrasts irrespective of whether the complementation markers involved in the contrasts make up morphosyntactic systems (i.e. distributionally delimited sets of expressions), but base our analyses on two kinds of evidence for semantic contrasts: minimal pairs and distributional facts. Obviously, the first kind of evidence – viz. minimal pairs – is stronger as it allows to better separate the semantic contribution of the complementizer from properties of the context. For instance, minimal pairs such as (1) is one piece of evidence that the contrast between *that* and *if* in English semantically is a contrast between semantic transparency and uncertainty about the complement proposition (*that* allows the factivity of *see* to pass on to the complement, whereas *if* cancels it).

(1) a. *John didn't see* [***that*** *Mary arrived from London*].
b. *John didn't see* [***if*** *Mary arrived from London*].

Similarly, the distributional fact that in northern Basque dialects the complementizer *-en* is found after complement-taking predicates like 'seem' and 'think', while *bait-* is found with predicates like 'it is a pity' and 'it is strange', is one piece of evidence that the contrast between these two complementizers has to do with factivity: 'seem' and 'think' are non-factive, whereas 'it is a pity' and 'it is strange' are factive predicates. Likewise, the specialization of one complementizer to proposition complements and of another to state-of-affairs (henceforth, SoA) complements can be diagnosed based on the compatibility of these complementizers with different types of complement-taking predicates. Proposition complements are found with predicates such as the following: assertive and interrogative utterance predicates, declarative knowledge predicates, non-immediate perception predicates, factive emotive predicates, epistemic modal or evidential predicates, propositional attitude predicates. In contrast, SoA complements are found with predicates such as these: directive utterance predicates, 'knowing how' predicates, immediate perception predicates, non-epistemic modal predicates, intention predicates and manipulative predicates.

2.1 Frequent semantic types

2.1.1 Proposition vs. state-of-affairs

The distinction between proposition and SoA is used here as an equivalent of Lyons' (1977) distinction between "third-" and "second-order entity"[2], and roughly equivalent with Vendler's (1967) distinction between "fact" and "event". Propositions and SoAs are both meanings prototypically expressed by clauses, but propositions can be understood basically as truth-valued meaning units, whereas SoAs can be understood basically as non-truth-valued meaning units. In a functional-cognitive approach, Boye (2010; 2012) argues that the distinction should be understood more precisely as pertaining to the construal of Langackerian processes (sequentially scanned relations): propositions evoke concepts construed as having a (situational) referent, whereas SoAs evoke concepts not construed as having a referent. Thus, Boye (2010) argues that propositions are conceptually more complex than SoAs: they add reference to the more basic SoA meaning, as reflected in the fact that proposition markers tend crosslinguistically to be at least as morphologically complex as SoA markers (cf. Boye 2012: 277–291). In addition, there is a crosslinguistic unidirectional implicational tendency that if a language has finite (or more precisely, what Stassen 1985: 76–83 calls "balanced") SoA-designating complements, then it also has finite (or balanced) proposition-designating complements (Cristofaro 2003: 122). In other words, the expression of a contrast between SoA and proposition by means of a contrast between respectively finite and nonfinite (or more precisely, what Stassen 1985: 76–83 calls "deranked") complements is unexpected.

A number of criteria exist for distinguishing propositions from SoAs. One criterion is that only propositions can be modified epistemically (e.g. Boye 2010, 2012: 195–198). Based on this criterion, the complement in (2a) comes out as propositional, while that in (2b) must be analysed as expressing a SoA: only the (2a) complement allows an epistemic adverb that scopes over the whole complement proposition.

(2) a. *He told her that he would (probably) leave her.*
b. *She told him to (?probably) leave her.*

In addition, proposition and SoA complements can be distinguished on distributional grounds. For instance, manipulative, desiderative and non-epistemic modal predicates typically allow only SoA complements, while knowledge, utter-

2 Lyons (1977: 141–142, 668, 723–724) restricts the term *proposition* to third-order entities designated by declarative clauses.

ance, epistemic attitude, emotive attitude and perception predicates allow of both proposition and SoA complements (Boye 2012: 188–194; see also Cristofaro 2003: 122). Thus, there is a contrast between telling and asking about propositions, as in (2a) and (3a), and telling and asking about SoAs, as in (2b) and (3b).

(3) a. *He asked her if she would leave him.*
b. *She asked him to leave her.*

Within this overall distribution pattern, there is some variation when individual complement-taking elements are considered, though. As one example of this, the Danish interrogative predicate *spørge* 'ask', unlike English *ask*, takes only balanced, propositional complements.

The distinction between proposition and SoA is crucial in order to understand clause semantics (accordingly, it plays a central role in Functional [Discourse] Grammar; Dik 1997; Hengeveld & Mackenzie 2008), and some of the distinctions discussed in subsequent sections of this paper may be seen as subdistinctions adding nuances to this basic one. For instance, the contrasts discussed in sections 2.1.2, 2.2.1 and 2.2.2 apply only to propositional complemens, while the realis-irrealis distinction in the specfic sense of this distinction discussed in Section 2.2.5 applies only to SoA complements.

Accordingly, the proposition-SoA distinction is the most pervasive of the semantic complement distinctions documented in this volume. It may be hypothesized that all of the languages covered code the distinction between propositional complements and SoA complemens. Not all contributions to this volume deal with the distinction, of course, but this is most likely due to the fact that proposition-SoA contrasts are not always coded in complementizer contrasts, the main focus of this volume. They seem to be more often coded in contrasts between finite and nonfinite complements, and such contrasts are likely to be ignored in studies dedicated to complementizers. Tellingly, Balkan languages are known for being poor on nonfinite complements, and all Balkan languages have complementizer contrasts between proposition and SoA. Below we first discuss cases in which the proposition-SoA contrast is coded in a complementizer contrast. Subsequently, we turn to the more frequent cases in which the proposition-SoA contrast is coded in a finite-nonfinite contrast.

Clear cases of complementizer contrasts between proposition and SoA are found in Latvian, Ossetic, Polish, Romanian and Turkish, in Permian languages, and as mentioned, in all Balkan languages irrespective of genealogical affiliation. A less clear case is found in Kalmyk, where the complementizer *giʁäd* seems to be restricted to SoA complements such as those of manipulative predicates, while *giǯə* applies to both SoA and proposition complements. Another less clear case is found in Maltese (see below).

In Turkish, the proposition-SoA contrast is coded in the contrast between *-DIK* and *-mA*. As discussed in Karakoç & Herkenrath (this volume), Csató (2010) proposes to analyse *-DIK* complements as designating "a proposition with a possible truth value", and *-mA* complements as designating "a predication without any possible truth value" (Csató 2010: 66), but of course it remains disputed whether *-DIK* and *-mA* are actually complementizers in a strict sense (see Karakoç & Herkenrath, this volume, for discussion). In Latvian, the proposition-SoA contrast is coded in the basic complementizer contrast between *ka* (proposition marker) and *lai* (SoA marker) (Holvoet, this volume). In (4a) the complement of 'know' designates a proposition, a piece of information which can be true or false, while in (4b) the complement of 'want' designates a SoA, a description of an action which can be carried out.

Latvian

(4) a. *Es* *zinu,* [***ka*** *tu* *atbrauksi* *paciemoties*].
1SG.NOM know.PRS.1SG **COMP** 2SG.NOM come.FUT.2SG visit.INF
'I know you will come on a visit.'
(Holvoet, this volume)

b. *Es* *gribu,* [***lai*** *tu* *atbrauc* *paciemoties*].
1SG.NOM want.PRS.1SG **COMP** 2SG.NOM come.PRS.2SG visit.INF
'I want you to come on a visit.'
(Holvoet, this volume)

Instead of *lai* Latvian employs the complementizer *kā* to mark SoA complements after some predicates. In (5a) and (6a) the complements introduced by *ka* are propositional, while in (5b) and (6b) they designate SoAs (see Sørensen & Boye 2015 on knowledge-predicate complementation, and Dik & Hengeveld 1991 and Boye 2010 on perception-predicate complementation).

Latvian

(5) a. *Viņš* *zināja,* [***ka*** *viņa* *spēlē* *klavieres*].
he know.PST.3 **COMP** she play.PRS.3 piano.ACC[PL]
'He knew that she played the piano'.
(Axel Holvoet, p.c.)

b. *Viņš* *zināja,* [***kā*** *spēlēt* *klavieres*].
he know.PST.3 **COMP** play.INF piano.ACC[PL]
'He knew how to play the piano'.
(Axel Holvoet, p.c.)

(6) a. *Viņš redzēja, [**ka** viņa spēlē klavieres].*
he see.PST.3 **COMP** she play.PRS.3 piano.ACC[PL]
'He saw that she played the piano'.
(Axel Holvoet, p.c.)

b. *Viņš redzēja, [**kā** viņa spēlē klavieres].*
he see.PST.3 **COMP** she play.PRS.3 piano.ACC[PL]
'He saw how she played the piano.'
(Nicole Nau, p.c.)

The SoA complementizer *kā* derives from a word meaning 'how'. Interestingly, several languages employ 'how' elements to introduce SoA complements. Like Latvian *kā*, the Polish complementizer *jak*, the Bulgarian complementizer *kak* and the Russian complementizer *kak*, all found in perception-predicate complementation, introduce SoA complements (see Hansen, Letuchiy & Błaszczyk, this volume). The same thing goes for Ossetic *k^{W}əd* (see Serdobolskaya, this volume, on Ossetic) and the 'how' complementizer used with perception predicates in Permian languages (Klumpp, this volume).

Romanian stands out among the Romance languages in that it seems to have a dedicated SoA complementizer, *să* (Fagard, Pietrandrea & Glikman, this volume). This complementizer enters into a three-fold contrast with the complementizers *că* and *dacă*, both of which are propositional and display an epistemic contrast of the type discussed in Section 2.1.2 below. A similar pattern is perhaps found in Maltese, where the SoA complementizer *biex* seems to contrast with two propositional ones, *li* and *jekk*. Other Romance languages have similar two-fold epistemic complementizer contrasts, but differ from Romanian in that their SoA complements are always nonfinite.

As mentioned above, distinguishing between finite and nonfinite complement is the most frequent way of coding the contrast between proposition and SoA in European languages: propositions are most often expressed by finite (indicative) complements, and SoAs by nonfinite ones (as already mentioned, it is telling that Balkan languages, which are poor on nonfinite complements, all have complementizer contrasts between proposition and SoA). Clear examples of this are found in Celtic, Finnic, Germanic, Permian and Romance languages (except Romanian), and in Hinukh, Hungarian and Ossetic. For instance, the Irish counterpart of the Latvian perception-predicate complement contrast in (6) above is a contrast between a finite complement (7a) and a nonfinite one (7b).

Irish

(7) a. *Chonaic mé* [***go*** *raibh sé ag damhsa*]
see.PST 1SG **COMP** be.PST.DEP 3SG.M PROG dance.VN
'I saw that he was dancing'.
(McQuillan, this volume)

b. *Chonaic mé* [*ag damhsa*] *é*
see.PST 1SG PROG dance.VN 3SG.M
'I saw him dancing'.
(McQuillan, this volume)

This is not to say that nonfinite complements always and exclusively designate SoAs. Boye (2010) argues that in Danish nominative- and accusative with-infitive complements, presence of the infinitival marker *at* ('to') marks the complement as propositional (8a), while absence of *at*, i.e. the occurrence of the morphological infinitive alone, marks it as SoA designating (8b).

Danish

(8) a. *Nationalbanken ses* [***at*** *stå for en ganske betydelig*
National.Bank.DEF see.PRS.PASS **to** stand.INF for INDF quite substantial
del af omsætningen].
part of trade.DEF
'The National Bank is seen to be responsible for a quite substantial part of the trade'.
(Boye 2010: 398)

b. *Bjarne Riis ses* [***Ø*** *stå og snakke med Sarevok*] ...
Bjarne Riis see.PRS.PASS stand.INF and talk.INF with Sarevok
'Bjarne Riis is seen talking with Sarevok'.
(Boye 2010: 398)

Also, finite complements may allow both a proposition and a SoA reading. This seems to be the case in Adyghe, where tensed, "factive" complements are used both to express propositional fact complements and SoA-designating complements of manner predicates. This may be a natural consequence of propositions being more complex than and involving SoAs (Boye 2012).

The identification of two patterns for expressing the proposition-SoA contrast – 1) complementizer vs. complementizer, and 2) finite (indicative) vs. nonfinite – should not be understood as suggesting that any given language exclusively exhibits one of these patterns, or that one and the same grammatical contrast is used in all the contexts in which it would be semantically relevant. Rather, several distinct grammatical contrasts may be employed, each with a limited distribution. For instance, the Danish contrast in (8) is used only with active appearance predicates ('seem', 'appear') and passive utterance, perception and propositional attitude predicates. With active perception predicates, Danish

uses a contrast between finite (indicative) complements and bare infinive complements to express a semantic contrast similar to that in (7). Similarly, Permian languages have two types of complements of perception predicates expressing immediate perception: 1) finite complements which involve a 'how' complementizer and specifically designate SoAs, and 2) nonfinite complements which employ a perfect participle and allow both SoA and propositional readings. As already noted, Latvian has two SoA complementizers, each with a distinct distribution, and Ossetic employs both dedicated complementizers and participles to mark SoA complements. Forker (this volume) identifies a number of distinct means of coding proposition-SoA contrasts in Hinukh.

It needs also be emphasized that the contrast is not always sharply cut. Firstly, rather than a strict distinction between either proposition or SoA, languages may make a distinction between proposition on the one hand and proposition or SoA on the other. Bulgarian and Russian *kak* complements can have either reading, and so can the Permian complements of perception predicates which involve perfect participles. Secondly, complementizers may be polyfunctional, having more functions than simply indicating propositions or SoAs. A case in point is the Lithuanian complementizer *kad*, which both marks SoA complements and a subset of propositional complements, contrasting with another propositional complementizer, *ar* (Holvoet, this volume). Thirdly, the binary proposition-SoA contrast may be part of a more complex contrast, as in the case of Romanian, where one SoA complementizer contrasts with two propositional ones that differ in terms of epistemic value.

We should like to round off the discussion of the proposition-SoA distinction by considering the two crosslinguistic tendencies pertaining to the distinction in the light of the data collected in this volume. As mentioned, there is a tendency for proposition markers to be at least as morphologically complex as SoA markers, and another tendency for the expression of embedded SoAs by means of finite complements to imply the expression of embedded propositions by means of finite ones. As far as we can tell, all the data collected in this volume conform to these two tendencies: we have encountered no contrasts in which a SoA marker is morphologically more complex than a proposition marker, and we have not come across complement contrasts in which the propositional counterpart was expressed by means of a nonfinite complement, while the SoA counterpart was expressed by means of a finite one.

We summarize the discussion above in Table 1 without pretending that the list of languages in the right column is all-inclusive.

Table 1: The coding of proposition-SoA complement contrasts in European languages.

Proposition-SoA contrasts coded in complementizer contrasts	Balkan languages irrespective of genetic affiliation, Latvian, Ossetic, Permian languages, Polish, Romanian, Turkish, (Maltese, Kalmyk)
Proposition-SoA contrasts coded in finite-nonfinite contrasts	Celtic languages, Finnic languages, Germanic languages, Hinukh, Hungarian, Ossetic, Permian, Romance languages except Romanian, non-Balkan Slavic languages (e.g. Polish, Russian)

2.1.2 Uncertainty vs. epistemic neutrality

European languages frequently have a complementizer contrast similar to the contrast between English *that* and *if* or *whether*. Several semantic analyses of these contrasts have been proposed (see Nordström & Boye, this volume, for discussion). For the purposes of this overview, we will adopt the analysis in Nordström & Boye (this volume; see Fagard, Pietrandrea & Glikman, this volume, for a distinct, but possibly compatible analysis). Accordingly, the contrasts discussed here are all analysed in terms of a distinction between uncertainty (complementizers corresponding to English *if* or *whether*) and epistemic neutrality or semantic transparency (complementizers corresponding to English *that*). Both epistemic uncertainty and epistemic neutrality pertains to propositions (see Boye 2012: Chapter 4 for detailed arguments for the link between epistemic meanings and propositions). Accordingly, complementizers that express these notions exclusively can in fact be thought of as proposition markers similarly to the propositional complementizers discussed in Section 2.1.1. They differ from those complementizers, however, in that they do not merely mark propositions, but are dedicated to indicating the epistemic evaluation of these propositions – or, in the case of epistemically neutral complementizers, to allowing the epistemic evaluation expressed or implied by the complement-taking element to pertain to the complement proposition.

We understand *uncertainty* complementizers as complementizers that express anything from lack of full certainty to complete uncertainty about the proposition expressed by the complement. Uncertainty complementizers can be recognized by the following typical features (Boye 2012):

1. They enter into harmonic combination with other uncertainty markers (e.g. adverbs such as English *perhaps*; see Boye 2012: 26 on the analysis of uncertainty and epistemic possibility as different construals of the same point of an epistemic scale). For instance, English *if* frequently combines harmonically with *perhaps*: *they will tell us if he is perhaps there*, but only rarely combines non-harmonically with *certainly*: *they will tell us if he is certainly there*.

2. They are functionally and often diachronically related to conditional protasis markers, as in the case of English *if* (see Section 5.1 for more details). The natural account of this is that also protasis markers can be analysed as expressing uncertainty: uncertainty about the cicumstances represented by the protasis proposition expressed by a circumstantial adverbial clause.
3. They are functionally and often diachronically related to polar interrogative markers. For instance, English *if* marks complements of the interrogative predicate *ask*, and in many languages counterparts of English *if* (as complementizer) are used in independent clauses as polar interrogative markes – cases in point are Irish *an*, Welsh *a*, Latvian *vai*, Lithuanian *ar*, Estonian *kas*, Komi *-ö* and Udmurt *-a* (see Section 5.1 for more details). The natural account of this is that polar questions imply uncertainty about the proposition they concern (see Boye 2012: 26–27, 308–315 for discussion).

With *epistemic neutrality* or *semantic transparency* we refer to complementizers that do not in themselves mean absence of any specific epistemic meaning, but rather are transparent in the sense that they allow the possible epistemic value of the complement-taking element to pertain to the complement proposition. The reason for considering such complementizers under the headline "epistemic" thus does not reside in the semantics of these complementizers, but in the fact that they contrast with other complementizers that do have a specific epistemic value, e.g. the value of uncertainty (cf. Boye, van Lier & Brink 2015). Transparent complementizers can be identified by the fact that since they do not have an epistemic value, they do not occur in harmonic combination with other epistemic markers, and thus – unlike uncertainty complementizers (see above) – do not show preferences for some epistemic markers over others. For instance, English *that* readily combines with both *perhaps* and *certainly*: *they told us that he is perhaps there*; *they told us that he is certainly there*.

In addition, uncertainty complementizers and epistemically neutral complementizers can be distinguished based on their distribution. Due to space considerations, we will refrain from a detailed discussion of this issue in this chapter and refer to the relevant chapters on individual language families. We would like, though, to draw attention to a pattern that seems to pertain more or less to all languages for which a complementizer contrast between uncertainty and epistemic neutrality can be identified: some complement-taking elements (notably interrogative predicates like English *ask*; see above) only allow an uncertainty complementizer; others (including propositional attitude predicates like English *belief*) only allow an epistemically neutral complementizer; still others (including knowledge predicates) allow both kinds of complementizers. In the latter

case, the choice between complementizers is used to make an epistemic distinction. Consider (9).

(9) a. *He knew* [*that she was there*].
b. *He knew* [*if she was there*].

In (9a), the epistemically neutral and transparent *that* allows the factivity of *know* to shine through to the complement proposition (outside metalinguistic, so-called "inverted commas" uses of *that*): truth of or certainty about the complement proposition is simultaneously presupposed and predicated of the subject referent of the main clause. In contrast, the complementizer *if* in (9b) expresses uncertainty about the complement proposition and thus blocks the factivity of *know*: knowledge is still predicated of the subject referent, but by means of *if* the speaker presents herself or himself as uncertain about the complement proposition.

Complementizer contrasts between epistemic neutrality and uncertainty are found in many European language families. Uncontroversial cases are found in Baltic, Basque, Celtic, Finnic, Germanic, Romance, Romani, Sami, and Slavic languages, and in Maltese, and Turkish. For instance, the Welsh contrast in (10) is similar to the English one in (9).

Welsh

(10) a. *Dw* *i* *'n* *gwybod* [***bo*** *hi* *fan hyn*]
be.PRS.1SG 1SG PROG know.VN **COMP** 3SG.F here
'I know that she is here.'
(McQuillan, this volume)
b. *Dw* *i* *'n* *gwybod* [***ydy*** *hi* *fan hyn*]
be.PRS.1SG 1SG PROG know.VN **COMP.be.PRS.3SG** 3SG.F here
'I know if she is here' (but other participants don't necessarily).
(McQuillan, this volume)

There is a considerable degree of variation in how precisely the contrast is expressed.

Firstly, in some languages the binary epistemic contrast is part of a more complex complementizer contrast. As discussed in Section 2.1.1, Maltese and Romanian show a triparte complementizer contrast between 1) proposition + epistemic neutrality, 2) proposition + uncertainty, and 3) SoA. Similarly, Latvian distinguishes not only between epistemic neutrality (*ka*) and uncertainty (*vai*), but also has a couple of dedicated SoA complementizers (*lai* and *kā*) (see Section 2.1.1). Somewhat differently, Germanic languages have two (or more) distinct uncertainty complementizers, e.g. English *if*, *whether*, German *ob*, *inwiefern*, Danish *om*, *hvorvidt* the choice between which has to do with a subtle meaning

difference (see Nordström & Boye, this volume, for discussion). In Maltese, *kemm* 'how much' can be used as an alternative to the canonical uncertainty complementizer *jekk*. Whereas *jekk* designates complete uncertainty, *kemm* in complements of a predicate of doubt indicates probability of the negative counterpart of the complement proposition (Borg & Fabri, this volume).

Secondly, the complementizer that covers epistemic neutrality may not do so unequivocally. In Lithuanian, *kad* is epistemically neutral, while *ar* expresses uncertainty.

Lithuanian

(11) a. *Jon-as* *ne-sakė,* [***kad*** *eis* *į* *koncert-ą*].
John-NOM NEG-say.PST.3 **COMP** go.FUT.3 to concert-ACC.SG
'John didn't say he would be going to a/the concert.'
(Holvoet, this volume)

b. *Jon-as* *ne-sakė,* [***ar*** *eis* *į* *koncert-ą*].
John-NOM NEG-say.PST.3 **COMP** go.FUT.3 to concert-ACC.SG
'John didn't say whether he would be going to a/the concert.'
(Holvoet, this volume)

However, *kad* is used also to introduce SoA complements. In (12), the distinction between proposition (12a) and SoA (12b) is expressed by means of, respectively, plain indicative and irrealis mood.

Lithuanian

(12) a. *Jon-as* *sakė,* [***kad*** ***pasitiks*** *mus* *stot-yje*].
John-NOM say.PST.3 **COMP** **meet.IND:FUT.3** 1PL.ACC station-LOC.SG
'John said he would meet us at the station.'
(Holvoet, this volume)

b. *Jon-as,* *sakė,* [***kad*** ***pasitiktume*** *jį* *stot-yje*].
John-NOM say.PST.3 **COMP** **meet:IRR.1PL** 3.ACC.SG.M station-LOC.SG
'John told us to meet him at the station.'
(Holvoet, this volume)

Thirdly, in some languages epistemic complementizers have a limited distribution. In Celtic languages, the contrast between epistemic neutrality and uncertainty is neutralized in negative complements. In affirmative complements, Welsh employs *y* to express epistemic neutrality and *a(i)* to express uncertainty, but in negative complements only the complementizer *na(d)* is used.

Welsh

(13) *Dw i 'n siwr* [***na*** *ddylech chi ddeud wrtho fe*]
be.PRS.1SG 1SG PRED sure **NEG.COMP** should.COND 2PL say.VN to.3SG.M 3SG.M
'I'm sure you shouldn't tell him.'
(McQuillan, this volume, referring to King 1993: 306)

In addition to or instead of an epistemic contrast of the type discussed above, some languages make a distinct type of contrast which may be seen as closely related. It seems that in Ukranian Servi Romani the contrast between the complementizers *kaj* and *so* is used for coding a contrast between different degrees of certainty about a complement proposition. In Russian, the epistemically neutral (or even "canonical"; see Hansen, Letuchiy & Błaszczyk, this volume) complementizer (*čto*) contrasts not only with an uncertainty complementizer (*li*), but also with a form composed (at least originally) by adding *-by* to the neutral complementizer: *čto-by*. Hansen, Letuchiy and Błaszczyk (this volume) analyse *-by* (which is usually glossed "conditional") as a subjunctive morpheme, and it seems to mark propositional complements as hypothetical rather than indicating a strictly epistemic value (degree of certainty, source of information), but in some cases the difference from epistemic complementizers may be seen as rather subtle. In the following example, from Polish rather than Russian, the natural interpretation seems to be clearly counterfactual, for instance (see Boye 2012: 28–29 on the analysis of counterfactuals as epistemic).

Polish

(14) *Czasami chcia-ł-by-m,* [***że-by-m*** *cię nigdy nie spotka-ł*].
sometimes want-PST-COND-1SG **COMP-COND-1SG** you.ACC never not meet-PST
'Sometimes I wished I never had met you.'
(Hansen, Letuchiy & Błaszczyk, this volume)

Permian languages do not seem to have a complementizer contrast of the specific type focused on in this section. However, at least Udmurt has a related contrast between epistemic neutrality marked by *šuysa* or absence of a complementizer, as in (15), and similative *budto(by)* 'as if'. As can be seen from (16), the latter word is distributed differently than English *as if*. For instance, it is found after propositional attitude predicates like 'think', and in this context at least, it can perhaps be analysed as epistemic (without necessarily being analysed as expressing uncertainty specifically).[3]

3 *Budto(by)* is borrowed from Russian. Neither for Russian nor for Permian languages is it entirely clear that it is indeed a complementizer.

Udmurt

(15) [*Boris araky-ez u-g jaraty kad*] *pot-e.*
Boris vodka-ACC NEG.PRS-3 love.CN as.if come.out-PRS.3SG
'It seems that (as if) Boris dislikes vodka.'
(Klumpp, this volume)

(16) *Aś-me-len tańi tros-ez malpa-lo,* [***budto*** *eksej, kalyk-ez*
self-1PL-GEN here many-3SG think-PRS.3PL **COMP:EPIST** tsar people-ACC
žaľa-sa syče mańifest pott-iz].
pity-CVB such manifest take.out-PST.3SG
'Many of us here think that the tsar may have issued the manifest in feeling compassion with the people.'
(Klumpp, this volume)

As mentioned, clear examples of complementizer contrasts between uncertainty and epistemic neutrality are found in Baltic, Basque, Celtic, Finnic, Germanic, Romance, Romani, Sami and Slavic languages, as well as in Maltese and Turkish. The fact that epistemic complementizer contrasts are so widespread in European languages is in line with crosslinguistic findings. Boye, van Lier & Brink (2015) found that 22 languages had epistemic complementizers (13 languages were excluded due to insufficient data). Among the types of epistemic complementizer contrasts distinguished, contrasts between epistemic neutrality and uncertainty were by far the most frequent (13 languages).

Based on their materials, Boye, van Lier & Brink proposed that these contrasts are subject to a crosslinguistic tendency for uncertainty complementizers to be at least as morphologically complex as epistemically neutral ones. Thus, they found contrasts in which both values are expressed by means of distinct overt complementizers, contrasts in which uncertainty is expressed by adding an extra element to the neutral complementizer, and contrasts in which epistemic neutrality is associated with the absence of the uncertainty complementizer, but they did not attest examples of contrasts in which uncertainty was associated with the absence of the neutral complementizer. In general our findings are in line with this tendency (see Section 3 on optionality), but Hinukh may be seen as providing a counterexample. In Hinukh, *-łi* is rougly an equivalent of English *that*, whereas uncertainty is associated with the absence of a complementizer. It should be noted though that the difference between presence and absence of *-łi* is not the only difference between epistemically neutral complements and uncertain ones (see Forker, this volume, for details).

The crosslinguistic frequency of epistemic complementizer contrasts suggests that such contrasts are important for language users. In the light of this suggestion, it is interesting to note that some languages that do not code the contrast

between epistemic neutrality and uncertainty about a complement proposition in distinct complementizers have other means for expressing the contrast. Kildin Saami exploits a complementizer distinction similar to that between English *that* and *if*, but in other Sami languages the contrast is expressed by means of distinct syntactic structures (Kotcheva & Rießler, this volume). In Udmurt, knowledge-predicate constructions corresponding to English *know + if* are formed by adding the "conditional" affix to the complement-taking predicate and employing a complement-type not marked for irrealis; in Komi, the same type of constructions have an "irrealis particle" in the independent clause and the "conditional" particle *kö* in the dependent one (Klumpp, this volume).

2.2 Less frequent semantic types

An important criterion applied in the selection of the semantic types presented in this section is that none of these types could be seen as a by-product of the basic semantic distinction between propositional and SoA complementation markers. Let us illustrate this with an example from Permian. Klumpp (this volume) argues that in Komi the choice between a finite complement clause with the clause initial complementizer *myj* and a deranked (participial) complement could be partly relegated to information structure. The participial complement is formally a possessive NP. The possessive relationship, Klumpp argues, prefers given information, and thus participial complements are inclined to convey given information, while new information is expressed by finite complements. In order to substantiate this claim, he provides an example showing that in elicitation his consultants accept participial complements to convey old information, while in contexts of new information they insist using a subordinate clause with canonical complementizer.

We suspect that this specialization of the participial construction to express old (given) information in Komi is related to its employment to describe SoAs. Likewise, the association of the finite (epistemically neutral) complementizer *myj* with new information is related to its propositional scope. The rationale behind this is that expressions rendering entities of higher conceptual complexity, such as propositions, that are independently tensed and epistemically evaluated, are recruited in the expression of the main informational input of the sentence; in other words, they serve to present new information. Conversely, expressions rendering less complex entities, like SoAs, which lack independent specification of tense and epistemic modality (an information redundant and recoverable from surrounding discourse; see Cristofaro 2007: 103–106 on information recoverability), are often recruited to convey given (old) information. We therefore assume

that the information structure contrast is not an independent variable in the functional make-up of the complementation markers in Komi, but is rather an emergent property of the differential marking of propositions and SoAs discussed in Section 2.1.1. Such semantic or pragmatic effects that could be derived from differences in the scopal properties of complementation markers are expelled from the following discussion.

The sub-sections below are organized according to the conceptual complexity of the complement: we start with semantic types pertaining to higher layer of complexity (propositions) and proceed with semantic types pertaining to lower layer of complexity (SoAs).

2.2.1 Factuality of propositions: factual vs. non-factual or neutral

We will reserve the term 'factive' (resp. 'non-factive') for the description of certain complement taking verbs, and will speak about 'factuality', where entire propositions are at stake. This subsection is concerned with situations where certain complementation markers are restricted to propositions with factual status.

In Basque, both methods for determining semantic functions of complementation markers – extracting meaning differences from minimal pairs of sentences, and studying the distribution of markers according to complement-taking predicates (henceforth, CTPs) – lead to substantial discoveries. The minimal pair of sentences in (17) provides evidence for the complementizer contrast between 'factual' and 'neutral'.

Western Basque

(17) a. *igarri/entzun dut* [*Jon etorri d-**ela***]
guess hear AUX John come AUX-**COMP**
'I guessed/heard that John came'
(Artiagoitia & Elordieta, this volume)

b. *igarri/entzun dut* [*Jon etorri d-**ena***]
guess hear AUX John come AUX-**COMP**
'I guessed/heard that John came'
(Artiagoitia & Elordieta, this volume)

The common Basque complementizer *-ela* occurs both after factive CTPs (like 'it is an advantage [that]') and non-factive CTPs (like 'believe', 'say', 'think', 'seem') (Artiagoitia & Elordieta, this volume), and thus could be considered to be epistemically neutral or semantically transparent (cf. Section 2.1.2). In Western Basque, this complementizer enters in factuality opposition with the complementizer *-ena*, which is licensed in contexts, where "the speaker generally takes for

granted the truth value of the proposition expressed by the complement clause" (Artiagoitia & Elordieta, this volume). For those speakers who use both *-ena* and *-ela* after verbs like 'know', 'guess' and 'hear' (as is the case in [17]), there seems to be a difference in the interpretation of the complement clause: "the choice for *-ena* leads to presuppose the truth of the complement clause, whereas *-ela* sentences [...] do not prompt any nuance on the truth value of the complement proposition". Artiagoitia and Elordieta back up their claim by showing that the complementizer *-ena*, which always presupposes the truth of the proposition, is incompatible with information focus interpretation of the complement clause, because informational focus provides new (non-presupposed) information (Artiagoitia & Elordieta, this volume). It follows that *-ena* is a factual complementizer and *-ela* is a non-factual (or semantically transparent) complementizer.

An example of semantic functions of complementizers defined purely on the basis of CTP distribution comes from the Lapurdian-Navarrese and Zuberoan dialects of Basque. In these dialects, the complementizer *-en* seems to be restricted to CTPs, like 'seem', 'think', 'be possible', which do not entail the truth of the complement proposition; in other words this complementizer seems to co-occur with non-factual complements only. In contrast, the prefixal complementizer *bait-* is restricted in the same dialects to CTPs, like 'it is a pity', 'it is strange', which entail the truth of the complement proposition (Artiagoitia & Elordieta, this volume). Based on these facts, *bait-* could be said to be factual complementizer, while *-en* is a non-factual complementizer.

Greek makes use of the factual complementizer *pu*, which is opposed to *oti*, introducing non-factual complements (Joseph, this volume). Consider (18) from Roussou (2010: 16, 19), which shows that these two finite complementizers may occur after the same CTP.

Greek

(18) a. *O Janis anisixi* [***pu*** *efijes*].
the John worry.3SG **COMP** left.2SG
'John worries about the fact that you left.'
(Roussou 2010: 19)

b. *O Janis anisixi* [***oti*** *efijes*].
the John worry.3SG **COMP** left.2SG
'John worries that you (may have) left.'
(Roussou 2010: 19)

Roussou argues that the complement clause in (18) is factual only when *pu* is present: in this case the leaving of the second person subject is taken for granted as a background information. If *oti* is used, the interpretation is non-factual: the worry concerns potential fact.

Another apparent case of factual complementizer comes from Adyghe. According to Serdobolskaya (this volume, on Adyghe), the factive verb form with the complementation marking prefix *zere-* is used only when the complement proposition is presupposed true. This form usually occurs with verbs of knowledge and utterance, as well as with emotive predicates. In complements, where the propositional content is not presupposed, the verbal form in the adverbial case is used. Compare the following examples with verbs of utterance.

Adyghe

(19) *s-jane* *je-s-ʔʷa-ʁ-ep* [*ocenke* *dej*
1SG.PR-POSS+mother OBL-1SG.A-say-PST-NEG mark bad
*qə-**zere**-s-hə-ʁe-r*]
DIR-**FCT**-1SG.A-carry-PST-ABS
'I haven't told mother that I got a bad mark.' (The speaker did receive a bad mark.)
(Serdobolskaya, this volume, on Adyghe)

(20) [*njepe* *weš'x* *q-je.š'xə-n-**ew***] *a-ʔʷa-ʁ*
today rain DIR-rain-POT-**ADV** 3PL.A-say-PST
'They said it would be raining today' (The sentence describes a radio forecast. The speaker does not know if the forecast will turn out true or false).
(Serdobolskaya, this volume, on Adyghe)

The factuality of the form in *zere-* follows also from the fact that the truth of the complement clause cannot be denied if this form is used. In case of denial, the form with adverbial case has to be used; compare (21) with (19) and (20) above.

Adyghe

(21) [*a-r* *qə-s-e-wa-ʁ-**ew***] *zəč'e-m-jə* *a-r-jə-ʔʷa-ʁ*
DEM-ABS DIR-1SG.IO-OBL-hit-PST-**ADV** all-OBL-ADD 3PL.IO-OBL-3SG.A-say-PST
'He told everyone that he had hit me' (but this is not true).
(Serdobolskaya, this volume, on Adyghe)

We can, thus, conclude that applied to propositions, the form in *zere-* is factual, whereas the form with the adverbial case is non-factual.

In some cases, it is not the complementizer but the antecedent of the complement in the main clause that encodes the distinction between factual and non-factual proposition. In Russian and Ossetic the factuality of complement propositions is conveyed by the presence versus absence of correlative element in the main clause. In Russian, the presence of the correlative element *to* is compatible with factual contexts (Hansen, Letuchiy & Błaszczyk, this volume) whereas the structure without the correlative is neutral with respect to factuality. While both the presence and the absence of the correlative are compatible with the verb *znat'*

'know' and other factive verbs of knowledge, non-factive verbs like *dumat'* 'think' and *ščitat'* 'suppose' can only host structures without correlatives; compare (22a) and (22b).

Russian

(22) a. *Ja* *znaj-u* ***(t-o)*** [*čto* *ničego* *ne* *znaj-u*].
I.NOM know-PRS.3SG **that-ACC.SG.N** that nothing.SG.GEN NEG know-PRS.1SG
'I know that I know nothing.'
(Hansen, Letuchiy & Błaszczyk, this volume)

b. *Ja* *dumaj-u* ***(*t-o)*** [*čto* *ničego* *ne* *znaj-u*].
I.NOM know-PRS.3SG **that-ACC.SG.N** that nothing.SG.GEN NEG know-PRS.1SG
'I think that I know nothing.'
(Hansen, Letuchiy & Błaszczyk, this volume)

The situation in Ossetic is identical: certain complementizers occur with and without correlative pronouns. A semantic factor responsible for the presence (resp. absence) of a correlative pronoun is factuality. Factual complements require (or at least are strongly associated with) correlative pronouns, while non-factual complements do not require correlative pronouns (Serdobolskaya, this volume, on Ossetic). Compare (23) with the correlative *wəj* and (24) without it:

Ossetic

(23) *dɜ=* *mad-ɜn* *sɜ-wəl* *nɜ* *žaχt-aj,* [*dəwwɜ* *kɜj* *ra-jšt-aj*],
thy mother-DAT what-SPESPL NEG say-PST.2SG two COMP PRVB-get-PST.2SG
wəj*?*
that.DEM
'Why didn't you tell mother that you got a bad mark?'
(Serdobolskaya, this volume, on Ossetic)

(24) *radio-jɜ* *ra-zərd-t-oj,* [*rajšom* *važal* *kɜj* *wə-zɜn*].
radio-ABL PRVB-tell-PST.3PL tomorrow frost COMP be-FUT[3SG]
'They said on the radio that it will be freezing tomorrow.'
(Serdobolskaya, this volume, on Ossetic)

While the complement clause in (23) presents a fact ('*x* got a bad mark'), the forecasted weather conditions in (24) are not considered to be facts.

2.2.2 Evidentiality

Estonian and Finnish encode the distinction between indirect evidentiality (inference or hearsay) and direct sensory evidentiality by means of the formal contrast

between the epistemically neutral (or semantically transparent; see Section 2.1.2) complementizer (Estonian *et*, Finnish *että*) and a special similative complementizer with the lexical semantics 'as if'. In both cases the complement verb is in the conditional mood; cf. the examples in (25) from Estonian.

Estonian

(25) a. *Talle* *tundub,* [***et*** *Jaan* *tuleks,* *kui* *me* *talle* *maksaksime*].
s/he:ALL feel:PRS.3SG **that** Jaan come:COND.3SG if we him pay:COND:1PL
'S/he feels that Jaan would come if we paid him.'
(Kehayov, this volume)

b. *Talle* *tundub* [***nagu*** *Jaan* *tuleks*].
s/he:ALL feel:PRS.3SG **as.if** Jaan come:COND.3SG
'It seems to her/him as if Jaan is coming' (at the very moment of observation).
(Kehayov, this volume)

While (25a) expresses logical inference, (25b) expresses direct sensory evidentiality (where the sensation remains unspecified). Such minimal pairs can be constructed also with CTPs expressing specific senses (like 'hear' or 'see') and adducing further evidential specification: with an auditory verb, the sentence in (a) would receive hearsay reading, while the usual reading of (b) would be 'direct auditory evidence'; with a verb of visual perception (a) would encode logical inference, while the (b) would be interpreted as direct visual evidence.

Russian, Baltic and Permian have similative expressions similar to those found in Finnic, but their status as complementizers as well as their meaning (epistemic or evidential) are arguable (cf. Hansen, Letuchiy & Błaszczyk, this volume; Holvoet, this volume; and Klumpp, this volume, for discussion), and therefore we refrain from speculation about their semantic contribution to the complementizer systems of these languages.

While Finnic, and possibly other languages included in this volume, attest for complementizer contrast along the values indirect (inference or hearsay) vs. direct evidentiality, the complementation marker system of Hinukh cuts the semantic space of evidentiality in another way. This language encodes the privative opposition between reported and unmarked. After the verb 'remember' and after certain verbs expressing knowledge and acquisition of knowledge, a complement marked by the quotative enclitic *ƛen* expresses that the propositional content is acquired from oral report, while in a complement marked with the abstract suffix *-ɬi* the source of information is considered unclear or unimportant, and is therefore left unspecified (Forker, this volume); cf. (26).

Hinukh

(26) a. *dіž* *c'aɬ-iš* [*Musa* *Ø-aq'e-s=**ƛen***]
1SG.DAT get.to.know-PST Musa(I) I-come-PST=**QUOT**
'I got to know that Musa came (*when I saw his hat).'
(Forker, this volume)

b. *de* *hayło-qo* [*me* *aƛ-a-do* *Ø-aq'e-ƛ'os-~~ɬ~~**i***]
1SG.ERG 3SG.M.OBL-LOC 2SG village-in-DIR I-come-HAB-**ABST**
c'aɬ-er-an
get.to.know-CAUS-INTFUT
'I let him know that you will come to the village.'
(Forker, this volume)

2.2.3 Quotation

There is a continuum between direct speech and reported indirect speech, which represents the degree to which the personal, spatial and temporal deixis of the original utterance is preserved. Many languages have quotative markers, which fulfill also complementizing tasks. The structures found in the scope of such quotative markers look like, but often slightly deviate from direct speech proper.

The first example comes from Hinukh. This and most of the other Nakh-Daghestanian languages have quotative enclitics (or particles), most of which are historically related to a converb of the verb 'say' (Forker, this volume). The enclitics feature as complementation markers that do not occur only after utterance verbs, but also after various types of CTPs: propositional attitude verbs, verbs of knowledge and acquisition of knowledge, evaluative predicates etc. In reported speech contexts (after verbs like, 'say', 'tell' or 'ask'), the Hinukh enclitic *ƛen* (and the respective enclitics in the other languages) functions as a quotative proper, in other words, as a marker of direct speech. According to Forker (this volume), the enclitic accompanies the reported utterance in its original form: there is no shift in spatial, personal or temporal deixis and the verb form is the same as in the corresponding main clause. The only difference between quotes and ordinary main clauses presenting direct speech concerns restriction on reflexivization: the speaker and the addressee are often encoded with first and second personal pronouns in the quote; however, reflexive pronouns used in the quote can only refer to the speaker, never to the addressee; see (27).

Hinukh

(27) *exo-qo-r* *Malla* *Rasadan-i* *eƛi-yo* [***“zo***
herdsman-LOC-LAT Mullah Nasredin-ERG say-PRS **REFL.SG**
q’iloza-qo-s *Ø-ekir-o”=**ƛen***]
board.OBL.PL-AT-ABL I-unbind-IMP=**QUOT**
‘Mullah Nasredin says to the herdsman: “Unbind me!”’
(Forker, this volume)

Lithuanian (Holvoet, this volume) has two particles – *atseit* and *neva*, which in independent sentences have evidential and epistemic functions. In addition to this, these particles also function as a kind of quotative complementizers; see (28) with *neva*. They are used to introduce utterances reproduced in their original form, retaining all their original features, including deixis, original discourse markers etc. Holvoet (this volume) notes that they differ from proper quotation markers only “in that the utterance they refer to need not be quoted literally, but a free interpretation of it is given”.

Lithuanian

(28) *Jie* *man* *sakė,* [***neva*** *ko* *ne-duodi* *pasimatuoti*
3.NOM.PL.M 1SG.DAT say.PST.3 **neva** why NEG-give.PRS.2.SG try.on.INF
šalm-o], *nors* *kit-as* *jau* *buvo* *jį*
helmet-GEN.SG though other-NOM.SG.M already be.PST.3 3.ACC.SG.M
užsidėj-ęs.
put.on-PPA.NOM.SG.M
‘They said to me *neva* why don’t you allow us to try on the helmet, though another [of them] had already put it on.’
(Holvoet, this volume)

Further evidence for quotative complementizers comes from Adyghe. Serdobolskaya (this volume, on Adyghe) presents some rarer complementation marking strategies in this language, including the grammaticalized form of the verb of speech *ʔʷe-* ‘say’ with the additive particle *-jə*.

Adyghe

(29) [*“se* *prosto* *sə-b-de-semerqewə-ʁ* *nah-ç’e”* ***ə-ʔʷ-jə***]
I just 1SG.ABS-2SG.IO-COM-joke-PST than-INS **3SG.A-say-ADD**
qə-s-jə-ʔʷa-ʁ.
DIR-1SG.IO-3SG.A-say-PST
‘“I have only played a practical joke on you, nothing else”, he said to me.’
(Serdobolskaya, this volume, on Adyghe)

In the course of grammaticalization, this form has undergone semantic bleaching to the extent that it can be used to mark complements of its diachronic source, the CTP *ʔʷe-* 'say' (as in [29]), without giving rise to tautology.

Ossetic has a quotative complementizer with a very similar distribution (Serdobolskaya, this volume, on Ossetic). The so called "citation particle" *žзʁgз* is a grammaticalized participle-converb of the speech verb *žзʁən*, which is used as a complementizer with various classes of CTPs, including verbs of perception, utterance and propositional attitude. Just like *ə-ʔʷ-jə* in Adyghe, this form is grammaticalized to the degree that it can introduce complements of its diachronic source without giving rise to tautology. After manipulative predicates expressing orders and requests this complementizer functions as a typical quotation marker, introducing direct speech and preserving the illocutionary force (e.g. the imperative mood) and the deictic center of the original utterance.

The grammaticalization of nonfinite forms of verbs of speech into quotative complementizers is attested also in other languages spoken at the eastern and southeastern outskirts of the European continent. One of the major complementation markers in Kalmyk *giʁäd* is a perfective converb form of the verb *gi-* 'say' (Knyazev, this volume). Knyazev notes that the sentence in (30), which would usually employ *giǯə* (another major complementation marker originating in the imperfective converb form of 'say'), sounds acceptable to his consultants with *giʁäd*, if the first person possessive marker in *aavəm* refers to the subject of the matrix clause, not to the speaker.

Kalmyk

(30)	??*Ajsa*	[*udlgo*	*aavə-m*	*ir-xə*	***gi-ʁäd***]	*kelə-v.*
	Ajsa	soon	grandfather-POSS.1SG	come-PTCP.FUT	**say-CVB.PFV**	tell-PST
	'Ajsa said my grandpa will come soon.'					
	(Knyazev, this volume)					

The fact that the preservation of the original deixis of the direct speech constitutes a licensing condition for the use of *giʁäd* leads us to the conjecture that the latter functions as a quotative marker in (30).

In addition to simplex complementation markers, complementizer combinations may also have quotative function. Estonian illustrates how a combination of complementizers correlates with deictic shift respectively lack of deictic shift in person agreement (Kehayov, this volume). The examples in (31) contain different structures used to encode a question addressed to the speaker (the deictic center of the semantically subordinate clause) by a man called *Jaan*.

Estonian

(31) a. *Jaan küsis:* [***Kas*** *sa ei tulegi Tallinnasse?*]
Jaan ask:PST.3SG **PQ** you NEG come.CN:CL Tallinn:ILL
'Jaan asked: Aren't you ever going to come to Tallinn?'
b. *Jaan küsis,* [***kas*** *ma ei tulegi Tallinnasse?*]
Jaan ask:PST.3SG **PQ** I NEG come.CN:CL Tallinn:ILL
'Jaan asked, if I'm never going to come to Tallinn.'
c. *Jaan küsis,* [***et*** ***kas*** *sa / ma ei tulegi Tallinnasse?*]
Jaan ask:PST.3SG **that** **PQ** you I NEG come.CN:CL Tallinn:ILL
'Jaan asked [aren't you / am I] ever going to come to Tallinn.'
(Kehayov, this volume)

While (31a) presents a direct speech with subject in the second person, as produced by the author of the question, and (31b) contains indirect speech with deictic shift from second to first person in the complement clause, example (31c) presents a mixture of indirect and direct speech. This sentence contains two elements, which alone serve as complementizers – *et* (the semantically transparent complementizer) and *kas* (polar question word). Only this complementizer combination licenses both, the second and the first person pronoun in (31c). In other words, this construction with combined complementizers, labelled "indirect quotative" by Kehayov (this volume), allows for both – preserving the original deictic orientation of the quotation, and for deictic shift typical for indirect speech (replacement of the second person marking with first person marking). Similar constructional ambiguity is attested in Kalmyk. Knyazev (this volume) notes that the complementizer *gižə* allows for both – preserving the original deictic orientation of the utterance and for a deictic shift typical for indirect speech. In his example (21) translated in English 'His father said that Badma should give *me* the money', the first person indexical marker corresponding to *me* in English could refer both to the speaker and the father.

2.2.4 Time

We have decided to approach the question about temporal functions of complementation markers rather restrictively. The first reason for this decision is related to the precarious issue about the dedication of linguistic form to certain function(s), the second to the scope of complementation markers.

Nonfinite complementation structures are often marked for tense (or taxis), typical examples being the Turkic and Finno-Ugric participles and converbs. In Nogai, for example, after certain types of CTPs (e.g. knowledge and propositional attitude predicates), the complement predicates in the *GAn*-form are non-future,

while the complements in the *(A)yAK*-form are future (prospective) (cf. Karakoç & Herkenrath, this volume). Likewise, in Finnic, the participial complements are differentiated according to tense: non-past (e.g. Finnish *-vAn*) vs. past (e.g. Finnish *-neen*) (see Kehayov, this volume). We will dismiss with cases like these, because these deranked verb forms are not dedicated markers of complementation but rather components of a template, which as a whole identifies pieces of syntactic structure as complements. This, in turn, implies that time is not a feature of the complementation marker, but a feature of its component, which is assigned to it independently of the procedure of identifying the sequential unit as a complement.

The second group of cases to be excluded is exemplified by the Ossetic complementizer *sɜmɜj*, which seems to occur only with complements describing hypothetical events (SoAs) that are posterior to the context of the main predication (see Serdobolskaya, this volume, on Ossetic). Note, however, that this complementizer is compatible only with SoA-complements (e.g. complements of manipulative verbs like 'tell to', 'want', 'must', 'let' and 'agree' that are not independently tensed), and therefore it cannot be said to possess temporal specification on its own. For this reason, we have chosen to focus on complementation markers occurring in complements that confine propositions.

A possibly legitimate example of time specification encoded by means of complementizer-switch is discussed in Kehayov (this volume); see (32) from Estonian.

Estonian (Kehayov, this volume)

(32) a. *Sa* *saad* *teada* [***kui*** *ta* *tuleb* *linna*].
you get:PRS.2SG know:*A/DA* **when** s/he come:PRS.3SG town.ILL
'You will get to know when s/he comes to the town.'
(Kehayov, this volume)

b. *Sa* *saad* *teada,* [***et*** *ta* *tuleb* *linna*].
you get:PRS.2SG know:*A/DA* **that** s/he come:PRS.3SG town.ILL
'You will get to know that s/he is coming to the town.'
(Kehayov, this volume)

The dependent clause in (32a) can be said to function as object of the verb 'know'. Kehayov argues that the temporal subordinator *kui* in this example ascribes a specific temporal interpretation to the sentence: the fact of someone being informed about someone else's arrival in the town is construed as (roughly) coinciding with the arrival. In (32b), on the other hand, the fact of someone being informed about someone's arrival is not anchored in a specific point in time, and the temporal interpretation of the sentence is non-specific. In terms of relative tense, the

'know' situation in (32a) is simultaneous with the dependent situation, whereas in (32b) it is anterior in relation to the dependent situation.

The only language, which has dedicated complementizers modifying the tense of the complement proposition is Irish. This language has series of complementizers that are marked for polarity, focality, epistemic modality and tense. If none of the constituents of the complement are focalized, the affirmative non-past epistemically neutral complementizer is *go*, while its past equivalent is *gur*; see the minimal pair in (33). The respective affirmative non-past complementizer, which marks epistemic uncertainty, is *an*, while its past equivalent is *ar*; see the minimal pair in (34).

Irish

(33) a. *Síl-im* [***go*** *bhfuil* *cathaoir* *anseo*]
think-1SG **COMP** be.PRS.DEP chair here
'I think there is a chair here.'
(McQuillan, this volume)

b. *Rinne* *muid* *dearmad* [***gur*** *imigh* *siad*]
make.PST 1PL forgetting **COMP.PST** leave.PST 3PL
'We forgot that they had left.'
(McQuillan, this volume)

(34) a. *D'fhiafraigh* *sí* *díom* [***an*** *mbriseann* *sé* *go* *réidh*]
PST\ask 3SG.F of.1SG **COMP.Q** break.PRS 3SG.M ADV easy
'She asked me if it breaks easily.'
(McQuillan, this volume)

b. *D'fhiafraigh* *sí* *díom* [***ar*** *bhris* *sé* *go* *réidh*]
PST\ask 3SG.F of.1SG **COMP.Q.PST** PST\break 3SG.M ADV easy
'She asked me if it broke easily.'
(McQuillan, this volume)

Unlike in (32), the complementizers in these examples are sensitive to the tense of the complement proposition, not to the tense of the main clause.

2.2.5 (Ir)realis

In Section 2.2.1 above, we discussed the distinction between factual and non-factual propositions, which was encoded by complementizers, or other features of the complementation construction, such as correlative pronouns. In this section we discuss a similar semantic distinction, which differs from factuality in that it does not concern propositions, but SoAs (see Section 2.1.1). As we will see, SoAs can be specified by complementation markers for their relation to reality.

Latvian encodes the privative distinction between irrealis and unmarked SoAs within its complementizer system. Holvoet (this volume) notes that in addition to the basic SoA complementizer *lai*, Latvian has another complementizer, *kaut*, which is often used "if the object of an act of volition is viewed as unreal". The examples in (35) do not differ only with respect to their complementizers; there is also opposition of mood: the irrealis complementizer *kaut* in (35b) is used only with irrealis mood on the complement verb, and, as Holvoet explains, "the low expectation of actuation" is further highlighted by irrealis marking on the matrix verb.

Latvian

(35) a. *Es gribu,* [***lai*** *tu klausi man-iem padom-iem*].
I.NOM want.PRS.1SG **COMP** 2SG.NOM listen.PRS.2SG my-DAT.PL.M advice-DAT.PL
'I want you to listen to my advice.'
(Holvoet, this volume)

b. *Es vēlētos,* [***kaut*** *tu klausītu man-iem padom-iem*].
I.NOM wish.IRR **COMP** 2SG.NOM listen.IRR my-DAT.PL.M advice-DAT.PL
'I wish you would listen to my advice.'
(Holvoet, this volume)

The SoA complementizer *lai* allows for optional irrealis marking on the complement verb, whereas with *kaut* the irrealis marking is obligatory, which suggests that *kaut* is a genuine irrealis complementizer.

If we conceptualize (ir)reality as a scalar notion, we find that in complements of verbs of fearing the Russian complementation marker system encodes the distinction between remotedly possible threat and more real threat. The complementation marking structure consisting of the word *kak* 'how' (occurring also alone as a complementizer) and the conditional mood plus negation of the complement verb in (36a) conveys that "the content of the feared state of affairs is portrayed as only remotely possible" (Hansen, Letuchiy & Błaszczyk, this volume). In contrast, the complementation marking structure consisting of the general complementizer *čto* 'that' and the Conditional mood in (36b) conveys a real threat.

Russian

(36) a. *Boj-u-s',* [***kak*** ***by*** *na ėt-o* ***ne*** *uš-l-i*
fear-PRS.1SG-REFL **how** **COND** on this-SG.ACC **not** leave-PST-PL
vs-e ego den'g-i].
all-PL.NOM his money-PL.NOM
'I am afraid that all his money might go on this.'
(Hansen, Letuchiy & Błaszczyk, this volume)

b. *Boj-u-s',* [***čto*** *on* ***by*** *ne* *spravi-l-sja*].
fear-PRS.1SG-REFL **COMP** he.NOM **COND** not succeed-PST-REFL
'I am afraid he wouldn't be able to do it.'
(Hansen, Letuchiy & Błaszczyk, this volume)

2.2.6 Polarity

Negation interferes in the marking of complementation in different ways. The complementation marker systems of the languages studied here allow for distinguishing between a) cases where a negation marker participates in the encoding/identification of the complement clause as such; b) cases where the complementizer is a typical negative polarity item, restricted in its occurrence to negated complement taking predicates; c) cases where the complementizer incorporates the negation of the complement clause.

The first type, which is not particularly relevant, because it does not concern the semantic contribution of the compelementation marker, but only its formal properties, is quite common in the eastern part of the continent (the only instantiation from elsewhere seems to be Maltese). This type could be further broken in two subtypes. In the first, the complement clause expressing alternatives consists of one positive and one negative verb form; see the example from Turkish with the converb form *olup* of the verb 'be' and the negated participial form of the same verb *olmadığı* (Karakoç & Herkenrath, this volume).

Turkish

(37) *Ooó* *şey* *malzeme-sin-e* *bağlı,* [*acı* ***ol-up*** ***ol-ma-dığ-ı***]
IJ thing ingredient-POSS.3-DAT dependent hot **be-CV** **be-NEG-NMLZ-POSS.3**
'Ooh, hang on, it depends on the ingredients whether it's hot.'
(Karakoç & Herkenrath, this volume)

Komi has a similar complementation marking structure; the so called "negative alternative structure" (Klumpp, this volume), which consists of an affirmative form of the complement predicate immediately followed by an equivalent form (with regard to tense and person) of a special verb expressing negation or by a negated existential verb. A similar structure, with an affirmative and negated verb occurs also in Hinukh (Forker, this volume).

The second structural subtype of a) occurs mostly after verbs of fearing and worrying, and involves only one, negated verb form in the complement. Klumpp (this volume) argues in his discussion of Komi, that in this case the negation is motivated by the fact that the state-of-affairs expressed in the complement is undesirable. The canonical complementizer *myj* may be present or absent from

the complement clause; cf. (38a) and (38b) from Komi-Zyrian, which have the same meaning; crucial in this language is the presence of the optative particle *med(ym)*, which licenses the negation of the complement predicate.

Komi-Zyrian

(38) a. *Me* *pol-a*, [***med*** *ošk-ys* ***o-z*** *žö* *vötćy* *me*
I fear-PRS.1SG **COMP:OPT** bear-3SG **not.PRS-3** MOD follow.CN I
börśa].
after
(Klumpp, this volume)

b. *Me* *pol-a*, [*myj* ***med*** *ošk-ys* ***oz*** *žö* *vötćy*
I fear-PRS.1SG COMP **COMP:OPT** bear-3SG **not.PRS-3** MOD follow.CN
me *börśa*].
I after
'I'm afraid the bear will follow me' (lit. 'I'm afraid that the bear would not follow me').
(Klumpp, this volume)

Similar structures are attested in Slavic (Hansen, Letuchiy & Błaszczyk, this volume) and Baltic (Holvoet, this volume).

Predicates of fearing and worrying are semantically similar to polar questions in that they presuppose the existence of opposite alternatives: actualization vs. non-actualization of a SoA. Therefore, it is not surprising that this type of complementation structure is found in indirect polar questions; in Maltese, for instance, the complementizer *jekk* 'if' occurs in a structure where the complement verb carries the suffix *-x*, which in independent clauses expresses negation, but, according to Borg and Fabri (this volume) is employed here to convey uncertainty; see (39).

Maltese

(39) *Saqsie-ni* [***jekk*** *n-af-**x*** *x'* *inhi* *r-risposta*].
ask.3M.SG.PFV.SBJ-1SG.OBJ **if** 1SG.IPFV-know-**UNC** what COP DEF-reply
'He asked me whether I knew what the answer was.'
(Borg & Fabri, this volume)

The type b), an NPI-complementizer, is found only in Basque dialects (Artiagoitia & Elordieta, this volume). The distribution of the complementizer *-enik*, occurring in Western and Central Basque, is generally determined by a negated CTP, an inherently negative verb (like 'deny') or, less frequently, by matrix question, that is, in typical negative contexts (see e.g. van der Wouden 1997; Giannakidou 2001). Compare example (40a) with the general finite complementizer *-ela*, which is not sensitive to the polarity of the complement taking predicate, and example (40b)

with the finite complementizer *-enik* occurring after negated matrix predicate; *-enik* would not be possible in the first example.

Basque

(40) a. *sinesten dut* [*Jon etorri d-**ela***]
believe AUX John come AUX-**COMP**
'I believe that John came'
(Artiagoitia & Elordieta, this volume)

b. *Gure alabak ez du sinesten* [*lurra eguzkia-ren inguruan dabil-**enik***]
our daugther:ERG NEG AUX believe earth sun-GEN around walk-**COMP**
'Our daughter doesn't believe that the earth revolves around the sun.'
(Artiagoitia & Elordieta, this volume)

The third type c) concerns distinct complementizers, which apart from identifying a clause as a complement, express whether the complement is negated or not. This type is attested in Irish and Welsh (McQuillan, this volume). Both Celtic languages have series of affirmative and negative complementizers encoding the general complementizer type corresponding to English *that*. The negative complementizer incorporates the negation of the complement predicate, which means that the latter is not separately negated[4]. Consider the Irish examples in (41).

Irish

(41) a. *Síl-im* [***go*** *bhfuil cathaoir anseo*]
think-1SG **COMP** be.PRS.DEP chair here
'I think there is a chair here'
(McQuillan, this volume)

b. *Síl-im* [***nach*** *bhfuil cathaoir anseo*]
think-1SG **NEG.COMP** be.PRS.DEP chair here
'I think there isn't a chair here'
(McQuillan, this volume)

Let us recapitulate the intersection points of complementation markers and negation: in type a) the negation is a formal property of a complementation marking template; in type b) a special complementizer signals that the matrix predicate is negated, and in type c) a special complementizer signals that the complement is negated and incorporates the negative operator. Only types b) and c) are pertinent to the issue of semantics of complementation markers.

4 In some rare cases Welsh allows adding a separate negator, but this is optional (see McQuillan, this volume).

2.2.7 Information structure

Several contributions to this volume observe that the choice of complementation marker is related to features of information structure, like focality, topicality, and givenness. In most of the cases, however, these information structure values are not part of the conventionalized meaning of the complementation marker in the complement, but are a byproduct of the recruitment of what could be identified as a complementation marker (or parts of it) in other semantic domains; recall the Komi example from the beginning of Section 2.2.

In Turkic, where complementation is mainly conveyed by means of nonfinite verbal forms with specific argument coding patterns, the choice of complementation marker seems to correlate with information structure. The Nogai participial complementation markers in *-(A)yatḳan* and *-(A)yatïryan* have progressive meaning and are highly focal. In contrast, the copular complementation marking forms *-AdI eken* and *-(A)r eken*, which designate "a more general intraterminal view without a sharp focus on the ongoing situation", are not focal (see Karakoç & Herkenrath, this volume). This means only that focality is associated with the feature [± progressive], not that it is a conventional contribution of the respective complementation markers to the sentence.

The situation in Adyghe is different: the form in *zere-*, which otherwise seems to be restricted to factual complements (recall the discussion on factuality in Section 2.2.1.), can be used in non-factual complements in case these constitute the topic of the sentence; see (42).

Adyghe

(42) [*marine* *bzəʎfəʁe* *daḳʷə* ***zere***-*χʷə-š'tə-r*] *a-rə-ʁe-š'tə-n.*
Marina woman seamstress **FCT**-become-FUT-ABS DEM-PRED-PST-AUX-POT
'It is likely that Marina will become a seamstress.'
(Gerassimov & Lander 2008, quoted by Serdobolskaya, this volume, on Adyghe)

Serdobolskaya (this volume, on Adyghe) concludes that the choice of complementation marker in Adyghe is partly determined by the information structure of the sentence; moreover, the feature of topicality overrides factuality as a factor influencing this choice.

Adyghe provides another example of the relationship between the choice of complementation marker and information structure; in SoA complements two complementation marking devices – the potentialis form in the adverbial case and the potentialis form in the instrumental case – show uneven distribution with respect to information structure; see (43) and (44).

Adyghe

(43) *we* *nah* *w-jə-č'ase-r* [*təʁe-m* *z-je-b-ʁewə-**n-ew***
you more 2SG.PR-POSS-favourite-ABS sun-OBL REFL.ABS-OBL-2SG.A-tan-**POT-ADV**
(-č'e)*] *ar-a? –* *hawəmjə* [*zə-b-ʁe-psḳə-**n-ew***] *ar-a?*
(-INS) DEM:PRED-Q or REFL.ABS-2SG.A-CAUS-bathe-**POT-ADV** DEM:PRED-Q
'Do you prefer tanning? Or bathing?'
(Serdobolskaya, this volume, on Adyghe)

(44) *w-e-š'ənе* [*w-jə-zaqw-ew* *wə-ḳwe-**n-ç'e***] *ar-a?*
2SG.ABS-DYN-fear 2SG.PR-POSS-alone-ADV 2SG.ABS-go-**POT-INS** DEM:PRED-Q
'You're afraid of going alone, aren't you?'
(Serdobolskaya, this volume, on Adyghe)

Serdobolskaya argues that if the complement is focused, or an element within the complement is focused, the form with the adverbial case is selected; see (43), where the instrumental case is not acceptable. The instrumental case can be used if the question focus is on the CTP, as in (44); complements with the instrumental case tend to be unfocused.

Another factor with relevance for the information structure is the presence of a correlative element. We observed that the presence of correlative pronouns in Ossetic correlates with factual complements (recall the discussion in Section 2.2.1). In addition, correlatives are associated with sentence topics, and, as the example in (45) shows, topicality is a stronger selectional factor than factuality. Despite the fact that the co-occurrence of the complementizer *kɜj* with the correlative pronoun *wəj* is associated with contexts where the complement is presupposed true, non-factual complements with this complementizer do take correlative pronouns if they are in topical position.

Ossetic

(45) *зž* *nɜ* *fe-q^{w}əšt-on*, [*χəl* ***kɜj*** *kotː-oj*], ***wəj***, *зmɜ* *mɜ*
I NEG PRVB-hear-PST.1SG quarrel **COMP** do-PST.3PL **that.DEM** and me.GEN
nɜ *wərn-ə*.
NEG believe-PRS.3SG
(Zaur quarreled with his wife!) '– I haven't heard that they quarreled, and I don't believe it.'
(Serdobolskaya, this volume, on Ossetic)

Although the complement clause in (45) is considered to be false by the speaker, the correlative pronoun is used because the clause is topical.

In addition to correlative elements, also complementizer combinations may exert influence on the information structure. Here too, the example comes from Ossetic (Serdoboslkaya, this volume, on Ossetic). In this language the coordinat-

ing conjunction *ɜmɜ* functions also as complementizer occurring with utterance, propositional attitude, evaluative and emotive predicates, and predicates of fearing; cf. (46) with *ɜnqɜlən* 'think'.

Ossetic

(46) *ɜnqɜl* *dɜn*, [***ɜmɜ*** *je=mbal-mɜ* *a-səd-iš*].
think be.PRS.1SG **and** 3SG.POSS=friend-ALL PRVB-go-PST.3SG
(Where is your brother? – He's not here) 'I think he went to his friend.'
(Serdobolskaya, this volume, on Ossetic)

The epistemically neutral propositional complementizer in Ossetic *kɜj* 'that', which can be used with the same types of CTP as *ɜmɜ*, and some other, like predicates of knowledge, could be combined with *ɜmɜ*; cf (47a) with *kɜj* and (47b) with *ɜmɜ* and *kɜj*. The difference between these two sentences forming a minimal pair, is that in (b) the complement clause is focalized, while in (a) it is not.[5]

Ossetic

(47) a. *ažɜmɜt* *š-ražə* *iš*, [*rašt* ***kɜj*** *nɜ* *wəd*], *u-wəl*.
Azamat PRVB-agree EXST right **COMP** NEG be[PST.3SG] that.DEM-SPESPL
'Azamat agreed that he was wrong.'
(Serdobolskaya, this volume, on Ossetic)

b. *ažɜmɜt* *š-ražə* *iš*, [***ɜmɜ*** *rašt* ***kɜj*** *nɜ* *wəd*],
Azamat PRVB-agree EXST **and** right **COMP** NEG be[PST.3SG]
u-wəl.
that.DEM-SPESPL
'Azamat agreed that he was wrong.'
(Serdobolskaya, this volume, on Ossetic)

In complements of manipulative predicates (e.g. 'want', 'let', 'require', 'tell to') *ɜmɜ* co-occurs with the complementizer *k^{w}əd* 'how; in order to'. Just like in the example above, *ɜmɜ* and *k^{w}əd* can feature alone in such complements, but the effect of their co-occurrence is focal.

The most apparent case of conventionalized information structure contrast comes from Celtic. Irish and Welsh have two distinct series of complementizers, which are marked respectively for the values [+ focal] and [– focal]. According to McQuillan (this volume), the normal word order in Celtic is VSO, and therefore a focus complement is "a complement beginning with a fronted element which is not a finite verb." This means that the Celtic languages use complementizers to focalize constituents, not entire complement clauses. Consider the following

5 Serdobolskaya (this volume, on Ossetic) notes, however, that this combination of subordinators is used to focalize not only complement clauses but also other types of subordinate clauses.

examples from Welsh with the epistemically neutral complementizers *y* and *mai*, the first introducing non-focused complements and the second focused complements.

Welsh

(48) a. *Dw* *i* *'n* *gwy-bod* [***y*** *bydd* *hi* *fan hyn*]
be.PRS 1SG PROG know-VN **COMP** be.FUT.3SG 3SG.F here
'I know that she will be here'
(McQuillan, this volume)

b. *Dw* *i* *'n* *gwy-bod* [***mai*** *nhw* *sy* *'n hwyr*]
be.PRS 1SG PROG know-VN **COMP** 3PL REL.PRS late
'I know that they're late'
(McQuillan, this volume)

Consider now the examples in (49), also from Welsh, with the uncertainty complementizers (corresponding to the English 'if, whether'): the non-focus complementizer *a* and the focus complementizer *ai*.

Welsh

(49) a. *Dw* *i* *'n* *gwy-bod* [***a*** *fydd* *hi* *fan hyn*]
be.PRS 1SG PROG know-VN **COMP.Q** be.FUT.3SG 3SG.F here
'I know if she will be here.'
(McQuillan, this volume)

b. *O'n* *i* *ddim* *yn* *gwybod* [***ai*** *fo* *oedd* *o*]
be.IPFV.1SG 1SG NEG PROG know.VN **COMP.Q** 3SG.M be.IPFV.3SG 3SG.M
'I didn't know if it was him.'
(McQuillan, this volume)

2.2.8 Aspectual type: episodic vs. generic

In this section we discuss the distinction between complements with generic (or gnomic) reference (as in *Grandpa hated lying to him*) and complements with episodic reference (*I hated it how you lied to me*.). As 'episodic' and 'generic' do not only concern the feature of 'limitedness' in the flow of time, but also time reference, this semantic distinction is one of the points where aspect and tense intersect (see e.g. Dahl 1995). In this section we discuss complementation markers adducing specifications relevant to the boundaries of activities/states, rather than locating these activities/states along categories, such as 'past' or 'future'; therefore we consider this complementation marker contrast to be aspectual rather than temporal. It is self-evident that only forms of complement-taking

predicates, which are underspecified for the distinction between episodic and generic, and thus allow for both, could build minimal pairs with the respective readings. Typical types of predicates, which license both episodic and generic complements, are emotive predicates, such as 'like' or 'hate', evaluative predicates, such as 'nice', 'pleasant', and predicates of fearing.

In languages where the complementation is carried out by syntactic templates with deranked verb forms, certain forms show specialization to generic contexts. In Adyghe, for instance, emotive verbs, like 'love', 'like' and 'bother', require the case-marked form of the potentialis in generic complements, see (50), while in episodic complements of these verbs (e.g. *I liked how he danced yesterday*) only the so called "factive form" and the verbal noun can occur.

Adyghe

(50) *č'ale-m* [*qe-ŝʷe-**nə-**r*] *jə-č'as.*
boy-OBL DIR-dance-**POT**-ABS POSS-favourite
'The boy likes to dance.'
(Serdoboslkaya, this volume, on Adyghe)

Another option in complements with deranked verb forms is that the distinction at hand is expressed by means of differential argument realization of the same verb form. Forker (this volume) discusses a semantic distinction, which makes Bagvalal exceptional among the other Nakh-Daghestanian languages. The alternation between the masdar form (a type of verbal noun) preserving the valency frame of a finite clause and the masdar form with genitive dependents reflects the semantic distinction between referential reading and non-referential reading. This means, according to Forker, that if the SoA described by the complement has a spatio-temporal anchoring referring to a specific situation, all arguments of the masdar form preserve their normal case marking, see (51a); if, in contrast, the complement does not refer to a specific situation, but rather to a generic type of situation, then a genitive dependent of the masdar is preferred; see (51b).

Bagvalal

(51) a. *ʕali-la* *q'očan-č'i* [*Rasul-i-r* *e-w* *w-aX:a* *w-it'e:-r*]
Ali-DAT want-NEG Rasul-OBL-ERG LOG-I I-away I-send-MSD
'Ali did not like how Rasul sent him away.'
(Forker, this volume)
b. *ʕali-la* *q'oč-in-o:-b* [*ima-š:u-b* *keč'* *b-ihi-r*]
Ali-DAT want-IPFV-PTCP-III father-OBL-GEN song III-take-MSD
'Ali likes father's singing.'
(Forker, this volume)

Also canonical complementizers are employed to encode the distinction between generic and episodic situations; in such cases the generic meaning is usually reserved for a temporal and/or conditional subordinator. This is the case in Russian (Hansen, Letuchiy & Błaszczyk, this volume) and Ossetic. In the latter, after verbs of positive emotion (such as 'like'), the compelementizer *k^{W}əd* 'how' is employed if the context is episodic; see (52a) below. According to Serdobolskaya (this volume, on Ossetic), in such examples the positive emotion arises as a result of specific situation described by the complement clause. In generic contexts, on the other hand, the complementizer *k^{w}ə* 'if, when' is employed; see (52b) below:

Ossetic

(52) a. *mɜ=* *žɜrdɜ-mɜ* *sɜw-ə,* [*dɜ* ***k^{w}əd*** *žar-əš*].
my heart-ALL go-PRS.3SG thou **how** sing-PRS.2SG
'I like your singing.'
(Serdobolskaya, this volume, on Ossetic)

b. *ɜž* *warž-ən,* [*χud-gɜ* ***k^{w}ə*** *fɜ-kɜn-əš*], *wəj.*
I love-PRS.1SG laugh-CVBP **when** PRVB-do-PRS.2SG that.DEM
'I love it when you laugh.'
(Serdobolskaya, this volume, on Ossetic)

2.2.9 Agent control

The notion of control concerns here the degree of control an agent has over an event.

Kehayov (this volume) refers in his discussion of complementation marker semantics in Finnish to Sands (2000), who explains the difference between the nonfinite complementation markers in (53a) and (53b) in terms of volitional control.

Finnish

(53) a. *Harkitsin* *jäisinkö* *Babyloniin* *opettelemaan*
consider:PST:1SG remain:COND:1SG:PQ Babylon:ILL learn:*MA*.ILL
lukemaan *tähdet*.
read:*MA*.ILL star:PL.NOM
'I considered whether to remain in Babylon to learn to read the stars.'
(Sands 2000: 145; cf. Kehayov, this volume)

b. *Harkitsin* *jäisinkö* *Babyloniin* *opettelemaan*
consider:PST:1SG remain:COND:1SG:PQ Babylon:ILL learn:*MA*.ILL
tähtien ***lukemista***.
star:PL:GEN **read:ACNMLZ**
'I considered whether to remain in Babylon to learn the reading of the stars.'
(Sands 2000: 145; cf. Kehayov, this volume)

The illative form of the *MA*-infinitive of the verb 'to read' in (53a) implies that the subject remained in Babylon with the purpose to learn how to read the stars; the subject has control over the learning process. The action nominalization in (53b), on the other hand, implies that the subject learns how to read the stars less intentionally, more through absorption (Sands 2000: 145). Thus, the illative form of the *MA*-infinitive implies as a complementation marker higher degree of control than the action nominalization. Kehayov (this volume) explains this effect as a possible by-product of the distinction between generic and specific referents: the Finnic *MA*-infinitive has an implicit but specific actor argument, whereas the action nominal has a generic actor argument: (53b) could be paraphrased 'I stayed in Babylon to learn how *people* read the stars'.

Matras and Tenser (this volume) show for Romani that higher control of the agent of the main clause over the SoA expressed in the complement is likely to correlate with the use of a simplex complementizer, usually the inherited SoA complementizer *te*. By contrast, weaker control of the agent of the main clause over the complement SoA often correlates with the use of a complex form (complementizer combination). This is best illustrated by the comparison of logophoric (same-subject) and non-logophoric (different-subject) complements of desiderative predicates. Agent identity across the two clauses presupposes stronger control over the fulfillment of the wish expressed in the matrix clause; see (54) from Polish Xaladytka dialect of Romani with the simplex SoA complementizer *te*. Conversely, agent of volition not coinciding with the subject of the complement clause presupposes that this agent has weaker control over the fulfillment of her wish; see (55), from Bergitka dialect of Romani (also spoken in Poland) with a complementizer combination, consisting of the propositional complementizer *kaj* and the SoA complementizer *te*.

Romani

(54) *kamj-om* [***te*** *dža-ł* *khere*]
wanted-1SG **COMP** go-INF home
'I wanted to go home.'
(Matras & Tenser, this volume)

(55) *kam-av* [***kaj*** ***te*** *oddža-ł*]
want-SG **COMP** **COMP** leave-3SG.SBJV
'I want him to go away.'
(Matras & Tenser, this volume)

Also Polish and Russian use a more complex construction to express lack of control (Hansen, Letuchiy & Błaszczyk, this volume). In same-subject contexts, instead of using the infinitive construction as in (56a), speakers of Russian may

intentionally use a biclausal construction with the general complementizer *čto* and the conditional (irrealis) particle *by*, in order to mark that they lack control over the event described in the complement; see (56b).

Russian

(56) a. *Ja* *xoč-u* *okazat'sja* *nevinovn-ym.*
I want-PRS.1SG find.oneself:INF:REFL innocent-INS.SG
'I wish to prove innocent.'
(Hansen, Letuchiy & Błaszczyk, this volume)

b. *Ja* *xoč-u* *čto-by* *ja* *okaza-l-sja* *nevinovn-ym*
I want-PRS.1SG that-COND I.NOM find.oneself-PST-REFL innocent-INS.SG
'I wish I could prove innocent.'
(Hansen, Letuchiy & Błaszczyk, this volume)

Irish exemplifies a situation where in complements of what is formally an evaluative predicate, but functionally desiderative predicate, different degrees of agent control are encoded by manipulation of the complementizers corresponding to English *that* and *if*. McQuillan (this volume) shows that the CTP *ba maith* 'would be good', which is used as a desiderative predicate, takes both, the epistemically neutral complementizer *go* in (57a), and the conditional protasis marker *dá* in (57b)[6].

Irish

(57) a. *Ba* *mhaith* *leis* [***go*** *mbeadh* *a* *bhean* *ann*]
be.COND good with.3SG.M **COMP** be.COND POSS.3SG.M wife here
'He would like his wife to be there'
(McQuillan, this volume)

b. *Ba* *mhaith* *leis* [***dá*** *mbeadh* *a* *bhean* *ann*]
be.COND good with.3SG.M **if** be.COND POSS.3SG.M wife here
'He would like (it) if his wife were there'
(McQuillan, this volume)

The difference between (57a) and (57b) reflects the degree of control exercised by the subject of the main clause over the situation described in the complement clause; in the first example this degree of control is higher than in the second. The lower degree of control in (57b) is iconic with a lower degree of syntactic integration of the main and subordinate clause: while the order of the clauses in (57b) can be reversed, in (57a) it cannot.

6 The *dá*-clause passes the test for a complement, as it fills the subject argument slot of the main predicate (McQuillan, this volume).

Studying the examples above, one may have noticed that the value of weaker agent control is often associated with irrealis marking of the complement clause; consider the above examples from Romani and Russian, where the irrealis forms are glossed as subjunctive and conditional respectively. This mapping between weaker control (and the corresponding lower expectation of realization of the complement event) and irrealis is evidenced by the Latvian example (35) presented in Section 2.2.5. In addition to the complementizer contrast (*lai* vs. *kaut*), there is a privative opposition in mood: unlike *lai*, the irrealis complementizer *kaut* is used exclusively with irrealis mood and this corresponds to a "low expectation of actuation" of the content of the wish of the speaker (Holvoet, this volume).

2.2.10 Summary: less frequent semantic types

The independent attestation of the same semantic type in two or more genetically unrelated or only remotely related languages could be seen as a sign of typological salience. All semantic domains presented in this section are encoded in the complementation marking systems of at least two languages which are genealogically sufficiently distant to guarantee typological salience; see Table 2.

Considering the data in Table 2 one may come to the erroneous conclusion that Germanic and Romance, which in terms of number of languages are among the most prominent families in Europe, do not encode any of the presented semantic types. This is not the case: their absence from the table is a consequence of the fact that the contributions on Germanic and Romance (Boye & Nordström, and Fagard, Pietrandrea & Glikman, respectively) focus on the major semantic types: propositions vs. SoAs, and epistemic neutrality vs. uncertainty. In terms of language coverage, these contributions are the most comprehensive ones; this breadth and representativeness is achieved, however, at the expense of narrowing the focus of the study to the major types.

Table 2: Less frequent semantic types encoded by complementation markers.[7]

Semantic type	Attested values[8]	Languages
Factuality	factual, non-factual	Basque Greek Adyghe
Evidentiality	indirect, direct sensory, reported	Estonian, Finnish Hinukh
Quotative	quotation, indirect speech	Hinukh Lithuanian Adyghe Ossetic Kalmyk Estonian
Time	simultaneous, anterior (taxis) non-past, past	Estonian Irish
(Ir)realis	irrealis, realis (could be gradual)	Latvian Russian
Polarity	negative polarity item, complement negator	Basque Irish, Welsh
Information structure	topicality, focality	Adyghe Ossetic Irish, Welsh
Aspect	generic, episodic	Adyghe Bagvalal Russian Ossetic
Agent control	strong control, weak control	Finnish Romani Polish, Russian Irish Latvian

7 Cases where the presence/absence of a correlative element is crucial in the marking of the given semantic type are not included in the table, as they are considered only partly relevant to complementation marker semantics.

8 Unmarked (or neutral) members of oppositions, i.e. members of binary oppositions, characterized by absence of the differential semantic feature, are not included in the list of values.

3 Complementizer optionality

By complementizer optionality we refer to semantically neutral (or near-neutral) alternations between presence and absence of a complementizer in otherwise structurally identical complements. We do so without taking a stance as to whether in these cases the complementizer is in any sense omitted or inserted.

It follows from the definition just above that complementizer optionality is distinct from cases where the choice between presence and absence of a complementizer is semantically significant, as in (58), where presence of the complementizer *kua* marks the complement as propositional, whereas absence marks it as SoA-designating.

Tukang Besi

(58) a. *No-'ita-'e* [***Ø*** *no-kanalako* *te* *osimpu*].
3R-see-3OBJ 3R-steal CORE young.coconut
'She saw him stealing the coconut'.
(Donohue 1999: 403)

b. *No-'ita-'e* [***kua*** *no-kanalako* *te* *osimpu*].
3R-see-3OBJ **COMP** 3R-steal CORE young.coconut
'She saw that he had stolen the coconut'.
(Donohue 1999: 404)

Hinukh provides a borderline case. As described by Forker (this volume), the abstract suffix *-łi* is optional after knowledge predicates, but there is a tendency for presence of the suffix *-łi* to indicate certainty about the complement proposition, and for absence to indicate uncertainty. Accoding to the above considerations, to the extent that this alternation is semantically significant, it is not a case of complementizer optionality.

It also follows from the definition that alternations like that in (59) are not cases of complementizer optionality.

German

(59) a. *Er* *sagt* [*dass* *er* *kommt*].
he say.PRS COMP he come.PRS
'He says he will come'.

b. *Er* *sagt* [*er* *kommt*].
he say.PRS he come.PRS
'He says he will come tomorrow'.

While the complements involve identical strings, they are nevertheless structurally distinct: only (59a) has a preverbal adverbial slot, as can be seen in (60).

(60) a. *Er sagt* [*dass er Morgen kommt*].
he say.PRS COMP he tomorrow come.PRS
'He says he will come tomorrow'.
b. *Er sagt* [*er kommt Morgen*].
he say.PRS he come.PRS tomorrow
'He says he will come tomorrow'.
c. **Er sagt* [*er Morgen kommt*].
he say.PRS he tomorrow come.PRS
'He says he will come tomorrow'.

In the European languages dealt with in this volume, complementizer optionality presents a great deal of variation. In some languages, such as Turkish, Baltic and Romance, all complementizers are obligatory, with the possible exception of dialects and colloquial speech (see Holvoet, this volume, on optionality in colloquial Baltic speech varieties). In other languages, at least one complementizer is optional. This seems to be the case in Basque, Finnic languages, Hinukh, Hungarian, Romani, Slavic languages, Permian languages, Maltese, and Welsh. This difference may be partly due to differences in complementizer semantics. Only certain semantic types may be qualified for optionality (see below).

But this cannot be the whole explanation. Even cognates from the same language family may in some languages be optional and in others obligatory. One example of this is what Nordström & Boye (this volume) refer to as Type 1 complementizers in Germanic languages (i.e. cognates or counterparts of English *that* and German *dass*): in Dutch, Afrikaans, German and Frisian, Type 1complementizers are obligatory (dialects and colloquial speech may present exceptions); in Danish, English, Faroese, Norwegian, Swedish, Icelandic and Yidish, on the other hand, they are optional, with Icelandic presenting an intermediate case in the sense that complementizer optionality is heavily restricted (for one thing, it requires that the complement subject is pronominal). Similarly, while in Welsh both Type 1 and Type 2 complementizers (the latter corresponding to English *if*) are fully optional, in another Celtic language, Irish, Type 1 is obligatory, and Type 2 can be dispensed with only in colloquial speech (McQuillan, this volume; see also Hansen, Letuchiy & Błaszczyk, this volume, on Slavic languages).

Complementizer optionality raises two questions (see also Nordström & Boye, this volume):

1. What governs the choice between presence and absence of the complementizer?
2. What determines the optionality of the complementizer?

The first question has been addressed in a number of publications (e.g. Elsness 1984; Kaltenböck 2009; Shank & Cuyckens, 2010; Boye, van Lier & Brink 2015; see also Nordström & Boye, this volume). A number of – probably partly language-

specific – factors influence the choice between presence and absence. Nordström & Boye (this volume) argue that for Germanic languages most of these factors can be seen as instantiations of one of the following general factors: 1) complement-taking elements undergoing grammaticalization or being discursively secondary (in the sense of Boye & Harder 2012) in relation to the complement favour absence of the complementizer). 2) Disfluency (possibly as a result of a relatively high degree of complexity) favours presence of the complementizer (which can serve as a pause marker or conversational turn holder). 3) Contexts in which absence of the complementizer leads to – temporary or non-temporary – ambiguity favours presence of the complementizer (cf. the fact that in Germanic languages, Type 1 complementizers are optional only when the complement is not sentence-initial).

Turning now to non-Germanic languages, we find factors that are similar to those that play a role in Germanic languages, but also other factors. As for similar factors, in Udmurt, like in Germanic languages with optional complementizers, optionality is cancelled if the complement is sentence-initial (Klumpp, this volume). In Maltese (Borg & Fabri, this volume), the complementizer *li* can be absent only after predicates that are used "parenthetically", i.e. with grammaticalized or discursively secondary status, just like in the relevant Germanic languages.

As for factors not known to play a role in Germanic languages, in Finnic, otherwise optional complementizers are not optional if the complement-taking element is a noun (Kehayov, this volume). The Maltese complementizer *jekk* can be absent only in presence of a special intonation contour (Borg & Fabri, this volume). In Udmurt, the complementizer *šuysa* is obligatory in short complements (e.g. if the complement consists only of a verb), but can be dispensed with if the complement-taking element is a finite form of the verb *šuy-* 'say'. This verb is the diachronic origin of *šuysa*, and absence of the complementizer thus prevents a *figura etymologica* (Klumpp, this volume). A final example comes from Romani, where absence of an optional complementizer is more frequent after inherited modals than after borrowed modals, and more frequent with unadapted borrowed complement verbs than with adapted borrowed complement verbs (i.e. verbs with an adaptation marker) (Matras & Tenser, this volume).

The second question mentioned above – what determines the optionality of the complementizer? – has received less attention than the first one. Nordström & Boye (this volume) suggest that in Germanic languages, optionality of Type 1 complementizers (i.e. complementizers like English *that*) is motivated by the transparency of these complementizers: they do not contribute anything semantically to the complement, but simply allow the epistemic stance expressed or implied by the complement-taking element to apply to the complement. Accordingly, the contrast between absence and presence of Type 1 complementizers is not semantically significant.

Boye, van Lier & Brink (2015) corroborate this account, but the authors point to data suggesting that in addition to semantically transparent complementizers, also complementizers indicating certainty about the complement proposition are sometimes optional. This may be seen as complying with a crosslinguistic tendency for the epistemic value of certainty to be zero-coded, i.e. to be conveyed by the absence of any explicit epistemic expression (Boye 2012: 176, 313–315).

Turning now to the optional complementizers covered in this volume, the majority seem to be epistemically neutral and semantically transparent, in line with the findings of Nordström & Boye (this volume) and Boye, van Lier & Brink (2015). This goes for Finnic 'that' complementizers, for Slavic 'that' complementizers (Russian *čto,* Polish *że,* Bulgarian *če*), for Permian 'that' complementizers (Komi-Zyrian *myj*, Komi-Permyak *što*, Udmurt *šuysa*), for Lihuanian *kad* and Latvian *ka* in colloquial speech, and for Hinukh *-ƛi* and *=ƛen*, Hungarian *hogy*, Welsh *y*, and Maltese *li*, in addition to 'that' complementizers found in some Germanic languages. However, there are also deviant cases. In Russian, Romani, Maltese and Welsh, not only the transparent complementizers (Russian *čto*, Romani *kaj*, Maltese *li*, Welsh *y*) are optional. In Russian, also the polar complementizer *-li* can be left out (in indirect questions and after 'don't undestand', 'don't know'). In Romani, also the non-factual complementizer *te* can be left out. In Maltese, also *jekk* 'if' can be dispensed with, provided that there is a special intonation. In Welsh, likewise, also *a* 'if' can be dispensed with. Irish complementizers, finally, seem to run entirely counter to the generalizations discussed above. In Irish, the transparent Type 1 complementizer *go* 'that' is obligatory, while Type 2 *an* 'if' may be left out in colloquial speech.

With the exception of Irish, though, the European languages seem to conform to the following tendency: If a European language has an epistemically neutral complementizer and allows for complementizer omission, one of the omissible complementizers is the epistemically neutral one.

4 Complementizer combination

In this section we focus mostly on combinations of canonical complementizers. Different complementation marking templates with deranked verb forms are mutually exclusive; they cannot combine and therefore are outside the scope of the section.[9]

9 In some marginal cases, however, such deranked forms can be combined with lexical complementizers (see e.g. Forker, this volume, on Lezgi).

Complementizer combination is defined here as a co-occurrence in the complement structure of two or more elements, which can identify complements on their own. The focus of this section is on the semantic effects, although in the end we will present some eye-catching structural idiosyncracies of such combinations.

We must concede that the idea of complementzer combination goes against the fundamental premise in the study of clause-linking syntax that clauses have only one complementizer position (e.g. Bresnan 1970). We do not intend to challenge this idea, which is compatible with an analysis of the attested "combinations" as singular complex complementizers. We are solely interested in the (compositional resp. idiomatic) semantics of combinations of elements which alone are employed to identify clauses as complements.

Our working hypothesis, based on non-systematic observations, was that if two or more complementizers are combined, one of them is always semantically transparent (see Section 2.1.2 on semantic transparency). The vast majority of the attested combinations indeed involves semantically transparent complementizers; cf. (61) from Romanian, where the semantically transparent propositional complementizer *că* 'that' co-occurs with *cum* 'how', which occurs also alone in complementizer position.[10]

Romanian

(61) *Mi-a* *spus* ***cum*** ***că*** *ar*
me.DAT-have.AUX.3SG tell.PST.PTCP **how** **COMP** have.AUX.COND.3SG
fi *plecat*
be.AUX.INF leave.PST.PTCP
'(S)he told me that (s)he would be gone'
(Fagard, Pietrandrea & Glikman, this volume)

Nonetheless, our working hypothesis is counterfeited by some complementizer combinations consisting of elements, neither of which is semantically transparent when functioning as a complementizer. Such combinations, characteristic mostly for colloquial speech, involve uncertainty markers of the 'if, whether'-type and are attested in Celtic (Irish, Welsh), Germanic (English, Swedish, Danish) and Finnish. For instance, colloquial Irish combines two polar complementizers: *cé acu* 'which one of two' and *an* 'whether':

10 Some informal versions of English allow using *how* in propositional complements of utterance verbs too; cf. *She told me how she has been in France*. (see Legate 2010)

Irish

(62) *Níl* *a* *fhios* *agam* [***cé acu*** ***an*** *ól-ann* *sí*] *nó* *nach*
NEG.be.PRS its knowing at.1SG **COMP.Q** **Q** drink-PRS 3SG.F or NEG
nól-ann
drink-PRS
'I don't know if she drinks or not'
(McQuillan, this volume)

Swedish combines *huruvida* 'whether' and *om* 'if'.

Swedish

(63) *Enligt* *rätten* *gick* *det* *inte* *att* *avgöra* [***huruvida*** ***om*** *hon*
according.to court.DEF go.PST it NEG to determine **COMP** **COMP** she
fallit *eller* *blivit* *slagen*]. (*Göteborgsposten* 2004)
fall.PTCP or become.PTCP hit.PTCP
'According to the court, it was impossible to determine whether she had fallen or been hit'.
(Boye & Nordström, this volume)

The weakest typological statement, which saves the importance of semantic transparency in complementizer combination, would be: if a language allows for complementizer combination, at least one of the combinations in this language contains a semantically transparent complementizer. Howbeit, even this statement is violated by the European data. The Celtic languages accept only combinations of complementizers expressing uncertainty over alternatives (as in [62] above from Irish), which means that they lack combinations involving semantically transparent complementizers.

An inquiry into the semantics of complementizer combinations in relation to the semantics of the respective individual complementizers warrants the following classification.

I. Semantically insignificant combinations: two or more complementizers are combined and the resultant meaning of the sentence is identical with the meaning of a sentence in which only one of them occurs, other things being equal. Three types of insignificant combinations can be distinguished, based on semantic transparency and meaning overlap.[11]

11 The first two types of "insignificant" combinations are "harmonic" in the sense that the co-occurring complementizers have (partly) overlapping meanings. It could be argued that harmonic combination reinforces the common meaning, in which case the combination is not entirely insignificant. Nevertheless, we will classify such combinations as 'insignificant' and will reserve

a) Combination of two semantically transparent complementizers. Most of the occurrences of this sub-type concern structural reinforcement of an original semantically transparent complementizer with its borrowed equivalent. In Komi-Zyrian, the inherited semantically neutral complementizer *myj* co-occurs with the semantically transparent borrowed *yštö* (< Russian *čto* 'that').

Komi-Zyrian

(64) *Me i-g dumajt,* [***yštö-myj*** *zerm-as*]
I NEG.PST-1 think-CN **COMP** rain-FUT3SG
'I did not think that it will rain.'
(Klumpp, this volume)

Udmurt, the other Permian language, combines the inherited semantically transparent complementizer *šuysa* with the transparent *štö*, also borrowed from Russian; cf. (65). Note that Udmurt is originally a head-final language and the inherited complementizer takes clause-final position; the borrowed complementizer, on the other hand, occurs in the beginning of the complement clause, and as such constitutes a sign of restructuring to head-initial syntax (Klumpp, this volume).

Udmurt

(65) *So malpa-s,* [***što*** *kreśan so-ly görškok-yn zarńi vaj-em*
(s)he think-PST.3SG **COMP** peasant (s)he-DAT pot-INE gold bring-PRF
šuysa].
COMP
'He thought that the peasant has brought him a pot of gold.'
(Klumpp, this volume)

b) Combination of two non-transparent semantically overlapping complementizers. This subtype was already exemplified by (62) from Irish and (63) from Swedish. Another example comes from colloquial Finnish, which combines the protasis marker *jos* 'if' with the polar question clitic *=ko/=kö*; both elements occur alone as markers of complements presenting alternatives.

Finnish

(66) *En tiedä* [***jos=ko*** *hän tulee*].
NEG:1SG know.CN **if=PQ** s/he come:PRS.3SG
'I don't know if s/he is coming.'
(Korhonen 1993: 127)

---

the label 'significant' exclusively for combinations adding a novel component to the semantic or pragmatic structure of the sentence

c) Combination of two semantically not overlapping, but compatible complementizers. This is by far the most frequent kind of complementizer combination, constituting more than half of the attested combinations. In all cases one of the co-occurring complementizers is semantically transparent; see (67) from Danish with *om* 'if' and *at* 'that'.

Danish

(67) *Reno ved ikke* [***om*** ***at*** *han skal have ondt af onkel Arthur*].
Reno know.PRS NEG **COMP** **COMP** he shall.PRS have sorry of uncle Arthur
'Reno doesn't know if he should feel sorry for uncle Arthur'.
(Boye & Nordström, this volume)

II. Semantically significant complementizer combinations: two or more complementizers are combined and the resultant semantics is not identical with the semantics of a sentence in which either of them occur, other things being equal. Such significant combinations are rare and their semantic contribution concerns peripheral functions, not related to modality. One group of instantiations of this type concerns the feature of control. In many dialects of Romani the propositional epistemically neutral complementizer *kaj* is combined with the SoA complementizer *te* to signal lower degree of control over the situation described in the complement. An example from the Romani dialects spoken in Poland was presented in Section 2.2.9; another example, from East Slovak Romani dialects, is presented in (68): the addition of the canonical *kaj* 'that' to *te* signals weaker semantic integration between the matrix and the complement SoAs, which in this case means less resolute relation between the request and its fulfillment.

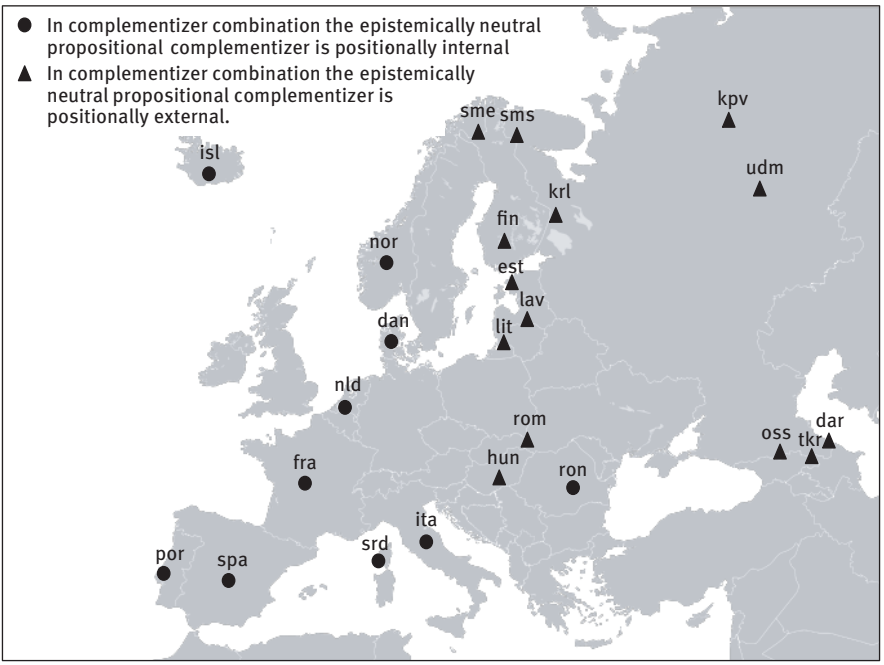
Romani

(68) *kam-av* [***kaj*** ***te*** *dža-l* *het*]
want-1SG **COMP** **COMP** go-3SG away
'I want him to go away'/'I wish that he goes away'.
(Matras & Tenser, this volume)

Other instantiations of this type concern quotations and information structure. As shown in Section 2.2.3, example (31c), the addition of the semantically transparent *et* 'that' to a question word functioning as a complementizer in Estonian produces construction which, unlike any simplex complementizer, allows for both preserving the original deixis of direct speech and for deictic shift typical for indirect speech. This ambiguity between quotative and reportative reading of the construction indicates that we are dealing with a novel form-to-function relation-

ship. An example of complementizer combination having effect on information structure was presented in Section 2.2.7, see example (47b). In Ossetic the coordinating conjunction *ɜmɜ* functioning also as a propositional complementizer was combined either with the epistemically neutral complementizer *kɜj* 'that' or with *k^{w}əd* 'how; in order that' to instigate focal complements.

Finally, let us briefly examine the relative order of the co-occurring complementizers. Most often we observe a combination of semantically transparent and semantically non-transparent canonical complementizers. This is the only combination type for which we have enough instantiations to dare typological generalizations and the only one showing a remarkable geographical pattern. The distribution on Map 1 prompts the following statements concerning the relative order of the semantically transparent (i.e. epistemically neutral) propositional complementizer and the semantically non-transparent complementizer in complementizer combinations.



Map 1: The order of semantically transparent and semantically non-transparent complementizers[12]

12 The language codes used in the map follow Ethnologue.

Statement 1: In the languages of eastern Europe in which this type of complementizer combination is attested, that is in Indo-European (Baltic, Romani [examples coming from Polish, Slovak, Ukrainian and Moldovan dialects], Ossetic), Finno-Ugric (Finnic, Sami, Permian, Hungarian) and Nakh-Daghestanian (Itsari Dargwa, Tsakhur), the semantically transparent propositional complementizer is always positionally external with respect to the non-transparent complementizer, viz. it is never found between the non-transparent complementizer and the core of the complement predication. In the head-initial (right-branching) languages this usually means that the semantically transparent complementizer occurs in clause initial position, while in the head-final (left-branching) languages (like Udmurt, Itsari Dargwa, Tsakhur) this usually means that the semantically transparent complementizer occurs in clause final position.

Statement 2: In the Germanic and Romance[13] languages in which this type of complementizer combination is attested, the semantically transparent complementizer is always positionally internal with respect to the non-transparent complementizer. In the light of the fact that Germanic and Romance languages are head-initial (right-branching), this means that the non-transparent complementizer occurs in clause initial position and is followed by the semantically transparent complementizer.

What is peculiar about the pattern on Map 1, is that despite the considerable mutability over time and the iconic nature of syntactic structucture, due to which it is rarely applied in historical reconstructions (cf. Salminen 2002), language families and their branches behave quite uniformly with respect to the relative order of elements in complementizer combinations, most of which, as we saw, are arbitrary in the sense that they lack conventionalized semantic functions.

5 Multifunctionality of complementizer forms

Linguistic forms which function as complementizers may have other grammatical functions. We distinguish two cases of multifunctional complementizer forms: a) cases where complementizer forms occur as subordinators in other types of dependent clauses, and b) cases where complementizer forms occur in independent clauses. We focus on canonical complementizers (in the specific sense rel-

13 Fagard, Pietrandrea & Glikman (this volume) speak about combinations of complementizers in the majority of the Romance languages, but the semantically non-transparent element functions in most of their examples as an adverbializer. We depart from their statements rather than from the examples provided, and therefore we consider the pairs of elements they mention as valid examples of complementizer combination.

evant here), as the contributions to this book contain too little information about the multifunctionality of entire complementation marking templates with nonfinite verb forms.[14] The discussion in this section is exclusively synchronic; we start by discussing the uses of complementizer forms as adverbializers and relativizers, and proceed by looking at the uses of complementizer forms in main clauses.

5.1 Complementizer forms are used as subordinators in other dependent clauses.

Let us first rule out the apparent possibility that all complementizer forms function as subordinators in other dependent clauses: this would entail that European languages lack dedicated complementizers. Although complementation is indeed most frequently encoded by means of multifunctional subordinators (which occur in adverbial and/or relative clauses), dedicated complementiers are attested.[15] Examples of complementizers which do not function as relativizers or adverbializers include the Polish epistemically neutral complementizer *że*, the Bulgarian polar complementizer *dali* 'whether' and its Romanian equivalent *dacă* (cf. Hansen, Letuchiy & Błaszczyk, this volume, and Fagard, Pietrandrea & Glikman, this volume), the Western Basque negative complementizer *-enik* and factual complementizer *-ena* (Artiagoitia & Elordieta, this volume), and the Hinukh quotative enclitic *=ƛen* (Forker, this volume).

There are about one hundred complementizers (or functional sets of complementizers, as for example, wh-markers) from different languages for which we have information about their employment as relativizers and adverbializers, or about their occurrence in independent main clauses. The contributions do not elucidate those issues systematically; one can suspect that many of the discussed complementation markers have adverbializer or relativizer uses, which are not mentioned in the respective contributions.

14 This lack of information is understandable, as the criteria for multifunctionality of syntactic configurations are far from being obvious. What concerns the nonfinite verb forms themselves, it should be also acknowledged that participles by definition function as attributes (or verbal adjectives), and as such represent relative (attributive) clauses, whereas converbs by definition function as verbal adverbs, and as such are representations of adverbial clauses (Ylikoski 2003). Hence, the question of whether nonfinite complementation markers, whose syntactic cores are morphological forms corresponding to the typological notions participle and converb could also function as relativizers and adverbializers, inevitably leads to circularity.

15 The question of whether the languages have complementizers restricted exclusively to complement clauses was not posed to the contributors; therefore, we refrain from judgments about the cross-linguistic frequency of dedicated complementizers.

Relativizers and adverbializers are defined here as subordinators, which introduce relative and adverbial clauses, respectively. Since there are various types of adverbial clauses, adverbializers are functionally more heterogeneous, and we will start with them.

Functionally, the prototypical class of complementizers is the epistemically neutral (propositional) complememententizers, like English *that*. The vast majority of items belonging to this class, altogether 27 complementizers from different languages and language families, are employed as adverbializers too.[16] Table 3 shows the occurrences of formal identity of epistemically neutral propositional complementizers and different types of adverbializers. Examples of forms, which function both as epistemically neutral propositional complementizer and as adverbializer introducing purposive clause include the Welsh *y*, Italian *che*, Danish *at* and North Saami *ahte*. Cases of identity of epistemically neutral propositional complementizers with markers of clauses of reason include Bulgarian *če* and Udmurt *šuysa*; with markers of clauses of time Basque *-ela* and Lithuanian *kad*, with markers of clauses of manner the Hungarian *hogy*, with markers of clauses of condition the Livvi-Karelian *ku*, with markers of clauses of result the Irish *go*, and with markers of clauses of comparison the Basque *-ela*.

Table 3: Formal identity of epistemically neutral propositional complementizers and adverbial subordinators

Polysemy according to the type of adverbial clause	Number of occurrences	Languages
neutral propositional – purpose	17	Irish, Welsh, Portuguese, Spanish, Italian, French, Danish, Latvian, Estonian, Finnish, Northern Karelian, Livvi-Karelian, Kildin Saami, Skolt Saami, North Saami, Udmurt, Adyghe
neutral propositional – reason/cause	9	Irish, Welsh, Bulgarian, Estonian, Livvi-Karelian, Kildin Saami, Udmurt, Komi, Old Zyrian
neutral propositional – time	6	Basque, Irish, Welsh, Lithuanian, Livvi-Karelian, Adyghe
neutral propositional – manner	3	Russian, Hungarian, Udmurt
neutral propositional – condition	2	Lithuanian, Livvi-Karelian
neutral propositional – result	2	Irish, Komi
neutral propositional – comparison	1	Basque

16 We refrain from generalizations about the identity of form between adverbializers and complementation markers with sub-propositional scope, as the latter are expressed in many languages of the sample by nonfinite constructions.

The data in the table suggests that the probability of different classes of adverbializers being expressed by the same forms as epistemically neutral propositional complementizers increases as follows (only significant intervals between the numbers in the table are included in the scale):

'comparison', 'result', 'condition', 'manner' → 'time' → 'reason' → 'purpose'

It is not easy to see the rationale behind this probabilistic scale. Epistemically neutral propositional complementizers typically occur after propositional attitude, knowledge and utterance predicates, and are associated with weak semantic connection between the content of the main and the complement clause (see Givón 1980; 2001: 40–41). On the other hand, cross-linguistically purpose clauses are often structurally identical with complement structures used after manipulative and desiderative predicates (Haspelmath 1989), which assume stronger semantic bond with their complements. Thus, the remarkable mapping between adverbial subordination of reason and purpose on the one hand and semantically transparent complementation marking, on the other, has to wait for explanation.

The semantic non-transparency of the other complementizers provides clues as to their adverbializer functions. We will discuss only the if/whether-type, which as shown in Section 2.1.2, encodes uncertainty, and which is the second most frequent and uniform finite complementizer type after the semantically transparent equivalents of the English *that*. The semantic notion of epistemic uncertainty occupies the middle point on the scale between full certainty in *p* and full certainty in *non-p*, therefore it is not surprising that it is carried on by expressions presenting alternatives (between *p* and *non-p*) (Boye & Nordström, this volume), such as *whether* (... *or*) and *if* (... *or*) in English. There are two types of complementation markers encoding this semantic type: i) markers occurring in complement clauses, but not in adverbial clauses, and ii) markers occurring in complement and adverbial clauses.

As a rule, Type I is synchronically or diachronically related to interrogative markers, whereas most of the elements in Type II originate in markers of conditional protasis. The employment of the complementizers of Type II type as temporal adverbializers always presupposes their synchronic use as conditional adverbializers. In addition, Estonian, Livvi-Karelian and some Germanic languages employ the uncertainty markers of Type II also as markers of adverbial clauses of comparison. This indicates that the distribution of Type II uncertainty complementizers in the adverbial domain is captured by the one-way implication: conditional ⊃ temporal/comparative.

Table 4: Formal identity of complementation markers encoding uncertainty and adverbializers

Type I: restricted to complement clauses	Type II: occurring in adverbial clauses	
	in conditional clauses	in conditional and temporal clauses
Irish *an*, Welsh *a* English *whether*, Icelandic *hvort*, Faroese *hvørt*, Norwegian *hvorvidt*, Swedish *huruvida*, Danish *hvorvidt*, Dutch *of* German *inwiefern, inwieweit*, Yiddish *tsi* Lithuanian *ar*, Latvian *vai* Russian *li*, Polish *czy*, Bulgarian *li*, *dali* Romanian *dacă* Northern Finnic (Finnish, Livvi-Karelian, Northern Karelian, Ludian) polar enclitics *=k(o)*/*=k(ö)*, Estonian *kas* Komi question enclitic *=ö*, Udmurt question enclitic *=a*	Irish *dá*, Irish *má* English *if*, Swedish *om*, Swedish *ifall*, Faroese *um*, Yiddish *oyb* Danish *hvis,* Icelandic *ef*, Dutch *als*, Spanish *si*, Portuguese *se*, Sardinian *si*, Italian *se*, French *si* Greek *an* Maltese *jekk* Latvian *ja* Russian *esli* Finnish *jos,* Northern Karelian *jos*, Ludian *d'eśl'i* Kildin Saami *koal'e, jesl'e*, Skolt Saami *jos*, Skolt Saami *jesl'i*, North Saami *jus* Hungarian *ha*	Estonian *kui*, Livvi-Karelian *ku* Ossetic *kʷə,* Ossetic *kɜd*

It can be seen from Table 4 that many languages have more than one complementizer encoding uncertainty. This is due to the two diachronic sources feeding this semantic type with linguistic matter: the interrogative source and the conditional source. This double origin and the presence of alternatives is probably explanatory as to question of why this semantic type is the one with most elements not occurring in other subordinate clauses: the availability of alternatives triggers specialization.

Summing up the discussion on formal identity of complementizing and adverbializing elements, we conclude that in Europe a random epistemically neutral propositional complementizer could be expected to have also a purposive function, while a random uncertainty complementizer could be expected to have also a conditional function.

Moving now to relative clauses, we observe that the identity of form of canonical complementizers and relativizers is somewhat less common than the identity of canonical complementizers and adverbializers; see Table 5. This holds true only if we focus on canonical subordinating conjunctions and exclude interrogative elements which occur both as relative pronouns and as a kind of "complementizers" taking an argument position in clauses often dubbed in the literature

"free (or headless) relative clauses" (e.g. *You can say what you please*.).[17] Here too, we can check only the epistemically neutral propositional type and the uncertainty type of complementizers for formal identity with relativizers, as these are the only complementizer types from which we have enough items in our corpus. We will also refrain from speculating about which type of relativizers, according to the kind of relative clause (restrictive, non-restrictive) they introduce, is more or less likely to display formal identity with complementizers, as we have too little data on this issue.

Table 5: Formal identity of complementizers and relativizers

Epistemically neutral propositional complementizers identical with relativizers	**Uncertainty complementizers identical with relativizers**
English *that* Romani *kaj* Maltese *li* Irish *go*, Irish *nach*, Welsh *y*, Welsh *na* Spanish *que*, Portuguese *que*, Sardinian *chi*, Italian *che*, French *que* Kildin Saami *šte*, Skolt Saami *što*, North Saami *ahte*	Basque *-en*

As can be seen from the table, the use of the epistemically neutral propositional complementizer forms as relativizers is quite common, although not as common as their use as adverbializers. Comparing Tables 3 and 5 we can see that while the epistemically neutral complementizer forms are employed as adverbializers in 24 languages from 11 language families (or upper level branches of families, such as Romance or Slavic), they are employed as relativizers only in 13 languages from 6 language families (or branches of families). Although the actual proportions probably deviate from these figures, as many of the contributions do not provide information about the use of neutral complementizers as relativizers, the observed difference is sufficient to claim that in Europe semantically transparent propositional complementizers are more often formally identical with adverbializers than with relativizers. It should be noted, on the other hand, that at least syncronically, the employment of epistemically neutral complementizer forms in relativization is independent of their employment in adverbialization. From the

17 Heine and Kuteva (2012: 168) argue that in the European languages such interrogative markers develop first to complementizers and then to relativizers.

items in the left column of Table 5, English *that*, Romani *kaj*, and Maltese *li* are employed as relativizers, but not as adverbializers.[18]

The right column of Table 5 shows that the use of complementation markers encoding epistemic uncertainty as relativizers is extremely rare, if not unique to Basque; cf. (69a) with *-en* as uncertainty complementizer and (69b) with *-en* as a relativizer (Artiagoitia & Elordieta, this volume). This could be explained by the fact that relative clauses provide background (presupposed) information about the entities they modify (Givón 2001: 176) and as such are incompatible with uncertainty.

Basque

(69) a. *zalantza* *daukat* [*Jon* *etorri* *d-**en***]
doubt have.1SG Jon come AUX-**COMP**
'I doubt whether John came.'
(Artiagoitia & Elordieta, this volume)

b. [*gure* *etxe-ra* *etorriko* *d-**en***] *lagun-a* *ez* *duzu* *ezagutzen*
our home-ALL come AUX-**COMP** friend-ART NEG AUX know
'You do not know the friend that will come to our house.'
(Artiagoitia & Elordieta, this volume)

5.2 Complementizer forms occurring in main clauses

We can distinguish between two types of main clause uses of forms that are employed in clausal subordination to introduce complements. In the first type the identical forms can be regarded synchronically as homonyms, despite their common diachronic source. The second type concerns cases where the function of the form occurring in main clauses is semantically contiguous (metonymical) to the complementizing function in subordinated clauses. In this case the diachronic split into complement and main clause uses is synchronically transparent. This type is captured by the notion of insubordination, as defined by Evans (2007).

In this section too, we will focus on canonical complementizers and will not treat nonfinite complementation markers.

18 This is in line with the cross-linguistically frequent path of development of complementation markers from relative strategies (Heine & Kuteva 2002: 325; Dixon 2006).

5.2.1 Homonymy

Forms that identify finite clauses as complements may represent various lexical classes when occurring in independent main clauses. Least uniform in this respect is the class of epistemically neutral complementizers, which display homophony with demonstrative pronouns (such as English *that*), with interrogative pronouns (such as Russian *čto* 'what', Komi *myj* 'what' and French *que* 'what') or proadverbs (such as Romani *kaj* 'where' and Hungarian *hogy* 'how'), with particles (such as Welsh affirmative particle *y*), with adpositions (such as Irish *go* 'to' and Welsh *i* 'to') and even with coordinating conjunctions (such as Irish *go* 'and' and Ossetic *ӕmӕ* 'and'). Complementizers encoding epistemic uncertainty, on the other hand, have more uniform main clause uses in our data: they mostly feature as polar question particles (e.g. the Irish *an*, Welsh *a*, Latvian *vai*, Lithuanian *ar*, Estonian *kas*, Komi *-ö* and Udmurt *-a*).

5.2.2 Insubordination

Nicholas Evans defined insubordination as a "conventionalised main clause use of what, on prima facie grounds, appear to be formally subordinate clauses" (Evans 2007: 367). There are various types of insubordination in the European data, but the most eye-catching regularity is that the scope of insubordinated complementizer forms in independent clauses tends to correspond to their complementizer scope.

Insubordinated complementation markers that were originally semantically transparent often come to express evaluations over propositions, be they positive, as in (70) with Hungarian *hogy* 'that', or negative, as in (71) with Lithuanian *kad*, which is used as a particle expressing skepticism and reservation.

Hungarian

(70) ***Hogy*** *Józsinak* *mekkora* *mázlija* *van!*
COMP Józsi:DAT how.big good.luck:POSS.3SG be[3SG]
'How lucky Józsi is!'
(Körtvély, this volume)

Lithuanian

(71) *Supranti,* *ką* *aš* *noriu* *dabar*
understand.PRS.2SG what.ACC 1SG.NOM want.PRS.1SG now
pasakyti? – ***Kad*** *gal* *ne-labai,* *dėd-e.*
say.INF **COMP** maybe not-much uncle-VOC
'Do you understand what I want to tell you now? – Well maybe not really, uncle.'
(Holvoet, this volume)

Epistemically neutral complementizer forms may also provide evidential qualification over the proposition in their scope; see (72) with Finnish *että* 'that' expressing report or inference.

Finnish

(72) ***Että*** *onkin* *kaunis* *paita!*
COMP be.PRS.3SG:CL beautiful.NOM shirt.NOM
'(And so) it is a beautiful shirt!'
(Kehayov, this volume)

Another frequent function of epistemically neutral or semantically transparent complementizers in insubordination contexts concerns deontic and directive contexts, in particular commands, requests, suggestions and permissions, as in (73) with French *que* 'that'.

French

(73) ***Qu'*** *il* *vienne!*
COMP he come.SBJV.3SG
'Let him come!'
(Fagard, Pietrandrea & Glikman, this volume)

The last semantic domain, which lends itself to insubordination with originally semantically transparent complementizers, is the domain of volition. This concerns mostly utterances expressing wishes, as in (74) with the Irish *go* 'that', or hopes, as in (75) with the Latvian *ka* 'that'.

Irish

(74) ***Go*** *gcuidí* *Dia* *leo*
COMP help.SBJV God 3PL
'May God help them!'
(McQuillan, this volume)

Latvian

(75) ***Ka*** *tik* *tante* *ne-pamana!*
COMP only aunt NEG-notice.PRS.3
'If only auntie doesn't notice!'
(Holvoet, this volume)

Complementizers co-occurring with SoA-complements, such as those found in the Balkan languages (Romanian *să*, South-Slavic *da*, Albanian *të*, Greek *na* and Romani *te*), are found in directives, expressions of deontic modality and wishes, but not in evaluative and evidential insubordinations. This restriction can be explained in terms of their scope properties: root modals address SoAs while evaluatives and evidentials address propositions (e.g. Boye 2012: 188–191, 195–198). These scope properties are directly reflected in the uses of complementizers in insubordination. Consider the examples with deontic modality and wishing in (76) from Bulgarian.

Bulgarian

(76) a. ***Da*** *na-piše-š* *tova* *pismo!*
COMP PFV-write-2SG.PRS this letter!
'You should really write this letter!'
(Joseph, this volume)
b. ***Da*** *s-te* *živi* *i* *zdravi!*
COMP be-2PL.PRS alive and healthy
'I wish you many happy years' (lit. 'May you be alive and healthy').
(Joseph, this volume)

The complementizer type encoding epistemic uncertainty often occurs in exclamative insubordinations marking wishes. In most of the cases this expressive function is carried out by complementizer forms, which mark conditional protases, such as the French *si* in (77). In some cases, complementizer forms of non-conditional origin are recruited: see (78) with the Romanian *dacă* 'if, whether', which has a compound origin (Fagard, Pietrandrea & Glikman, this volume)

French

(77) ***Si*** *seulement* *je* *pouvais* *lui* *manquer*
if only I can.PST.1SG 3SG.DAT miss.INF
'If only he could miss me.'
(Fagard, Pietrandrea & Glikman, this volume)

Romanian

(78) ***Dacă*** *(măcar)* *ar* *veni*
if only have.COND.3SG come.INF
'If only he could come.'
(Fagard, Pietrandrea & Glikman, this volume)

Just like in the case of homonymy, the functional range of the epistemically neutral propositional complementizers in insubordination shows fewer limits than the functional range of other complementizer types. This is also shown by the fact that the neutral type is more likely to manifest idiosyncratic extensions with restricted distribution. For instance, the use of the Italian *che* 'that' to mark a polar question in (79) does not seems to have parallels in the other Romance languages.

Italian

(79) ***Che*** *hai* *sonno?*
that have.2SG sleep
'Are you asleep?'
(Fagard, Pietrandrea & Glikman, this volume)

6 Diachrony of complementizers

Contemporary functionalist research on complementation has seeked to explain the syncronically observed distribution of different complementation markers across different complement taking predicates in terms of general principles, such as for example the iconicity between the degrees of semantic and syntactic integration of the matrix and complement clause (Givón 1980). Cristofaro (2014) contested this essentially ahistorical approach by showing that it fails to capture the distribution of a significant number of complementation structures from various languages. Complementation markers have developed along specific diachronic paths, which often lead to idiosynchractic uses, which go against models built on synchronic observations. Therefore, an examination of the diachronic processes involved in the development of complementation markers is indispensible part of the study of their semantic specialization.

In this section we will again dismiss with nonfinite constructional templates serving as complementation markers, as most of the papers do not contain information about these in their languages. We will concentrate on the diachronic sources of epistemically neutral and the uncertainty complementizers, as these

are the only semantic types for which we have enough diachronic data to dare generalizations.

6.1 The epistemically neutral type

6.1.1 Frequent sources

Epistemically neutral propositional complementizers exhibit quite similar sources in the European languages: in the vast majority of cases, the neutral complementizer originates in a proform (usually pronoun, sometimes proadverb), which in the course of grammaticalization has lost its constituent status. In most of the cases the ultimate source is either a demonstrative (anaphoric) pronoun, or an interrogative proform, which in many cases functions as a relativizer as well. The respective forms belonging to these two source classes are[19]:

a) demonstrative: Afrikaans *dat*, Danish *at*, Dutch *dat*, English *that*, Faroese *at*, Frisian *dat*, German *dass*, Icelandic *að*, Nowegian *at*, Swedish *att*, Finnish *että*, Estonian *et*, Northern Karelian *jotta*;
b) interrogative(-relative): Spanish *que*, Portuguese *que*, Sardinian *chi*, Italian *che*, French *que*, Latvian *ka*, Lithuanian *kad*, Greek *oti*, Russian *čto*, Polish *że*, Bulgarian *če*, Romani *kaj*, Veps *mi*, Livvi-Karelian *ku*, Komi *myj*. This seems to be the most salient diachronic source of neutral propositional complementizers in Europe.

Cases where the complementizer function arises through a relativizer stage are known from the literature on European languages (see Cristofaro 1998); typical examples in our sample are the Greek *oti*, Lithuanian *kad* and its Old Lithuanian equivalent *jog*.

6.1.2 Occasional sources

In addition to the recurrent sources presented above, we have two other cases where semantically neutral complementizers have identical sources in non-related languages.

In the Turkic contact area forms of utterance verbs (such as 'say') tend to grammaticalize into quotative complementizers; cases in point are Hinukh *-ƛen*

19 We exclude from the list cases where complementizers have been borrowed.

and its equivalents in most of the other Nakh-Daghestanian languages, Turkish *diye*, Nogai *dep*, Adyghe *ə-ʔʷ-jə*, and Ossetic *žзвgз*. Through desemantization and context generalization such complementation markers may later become semantically neutral; the examples from our sample are the Udmurt *šuysa* and the Kalmyk *gižə*, which are the semantically transparent complementizers in these languages.

The second source is represented by Irish *go* and Basque *-ela*, going back to a lative preposition ('to'), which has developed into manner adverbializer and later to neutral complementizer.

Peculiar sources of origin of neutral propositional complementizers, with only one occurrence in the European sample, are the Welsh *y*, which derives from a preverbal affirmative particle and the Welsh *mai* deriving from the third person singular present indicative form of the verb 'be'.

6.2 The uncertainty type

6.2.1 Frequent sources

In terms of diachronic sources the epistemic uncertainty type is even more uniform than the semantically neutral type. Virtually all conventionalized uncertainty complementizers may derive either from polar question markers or from conditional protasis markers. For the following complementizers there is either diachronic evidence that they developed from a polar question marker or a conditional protasis marker, or there is synchronically a clear formal relation to a polar question marker or a conditional protasis marker:

a) polar question marker (particle, clitic, pronoun) or a compound containing polar question marker: English *whether*, Danish *hvorvidt*, Faroese *hvørt*, Icelandic *hvort*, Norwegian *hvorvidt*, Swedish *huruvida*, Irish *an*, Welsh *a*, Finnish *=ko/=kö*, Ludian *=go*, Veps *=k*, Estonian *kas*, Russian *li*, Polish *czy*, Bulgarian *li*, Bulgarian *dali*, Latvian *vai*, Lithuanian *ar*;
b) conditional protasis marker: Spanish *si*, Portuguese *se*, Sardinian *si*, Italian *se*, French *si*, English *if*, Faroese *um*, Danish *om*, Norwegian *om*, Colloquial Icelandic *ef*, Swedish *om*, Colloquial Swedish *ifall*, Colloquial Dutch *als*, Yiddish *oyb*, Maltese *jekk*, Adyghe *-me*.

6.2.2 Occasional sources

The only complementizers encoding epistemic uncertainty, for which we know that they are neither related to the marking of polar alternatives nor to protasis marking, are the German compound items *inwiefern* and *inwieweit*.

6.3 Borrowing

Table 6 presents in absolute numbers the frequency of borrowing[20] of semantically neutral and uncertainty complementizers. The list of the borrowed complementizers includes:

a) borrowed semantically neutral complementizers: Finnish Romani *at* (< Swedish), Finnish Romani *et* (< Finnish), Xoraxani Romani dialect of Bulgaria *či* (< Bulgarian), Xoraxani Romani dialect of Bulgaria *ani* (< Gagauz/Turkish[21]), Romani of Romania *ke* (< Romanian), Lombard Sinti *ke* (< Italian), Romani from Hungary and Romania *hod ~ hodž ~ hod ~ hot ~ hoj* (< Hungarian), Romani dialects from Russia, Ukraine, Lithuania and Estonia *so* (< Russian), Romani of Greece *oti* (< Greek), North Saami *ahte* (< Finnish), Skolt Saami *että* (< Finnish), Skolt Saami *što* (< Russian), Kildin Saami *šte* (< Russian), Livvi-Karelian *što* (< Russian), Ludian *što* (< Russian), Komi-Permyak *što* (< Russian), Turkish *ki* (< Persian), Tsakhur *ki* (< Persian), Lezgi *x̂i* (< Persian), Kryts *ki* (< Persian);
b) borrowed uncertainty complementizers: Yiddish *tsi* (< Polish), Romani in Russia *li* (< Russian), Romani in Poland *či* (< Polish), Romani in the Balkans *dali* (< Balkan Slavic), Kildin Saami *jesl'e* (< Russian), Kildin Saami *koal'e* (< Russian), Skolt Saami *jesl'i* (< Russian), Skolt Saami *jõs ~ jos* (< Finnish), North Saami *juos ~ jus ~ jos* (< Finnish).

In all cases the complementizers are borrowed together with their complementizer semantics in the source language.

20 We subsume under borrowing only cases of matter replication, and no cases of pattern replication, although the latter are attested in the sample. For instance, Lotfitka Romani *sy* seems to be a calque from Latvian, which mirrors the homophony of the interrogative word 'how' and the semantically neutral complementizer in Latvian (Matras & Tenser, this volume).

21 The word *ani* is used as a complementizer in Gagauz but not in Standard Turkish (Birsel Karakoç, p.c.; Menz 1999: 67–69).

Table 6: Frequency of complementizer borrowing according to semantic type

Semantic type of complementizer	Frequency of borrowing		Diversity	
	Total of borrowed complementizers	Total of complementizers with determined historical origin	Total of languages with borrowed complementizers	Distinct families/ languages with borrowings
neutral propositional	20	59	10	Romani, Finnic, Sami, Permian, Turkic, Nakh-Daghestanian
epistemic uncertainty	9	50	5	Romani, Yiddish, Sami

The distribution in the table shows that semantically neutral propositional complementizers are more susceptible to borrowing than complementizers encoding epistemic uncertainty. Out of a total of 59 semantically neutral propositional complementizers with identified origin, more than one third are borrowings, whereas out of 50 complementizers of the uncertainty type, only 9 are borrowings. In order to minimize the distortion of the sample caused by Romani, where we have many borrowed items in the dialects of a single language, we counted the number of languages in which borrowing is attested (counting Romani as a single language). The result of this count is 10 languages from 6 distinct upper level branches of linguistic families, where borrowing of the semantically neutral type is attested, against 5 languages from three branches, where borrowing of the uncertainty type is attested. Thus, the greater disposition of the semantically transparent type to borrowing compared to the uncertainty type is confirmed based of three variables: number of instantiations of borrowing, number of languages with such instantiations, and genealogical distance between languages manifesting borrowing.

Matras (2009: 162, 196) and Matras and Tenser (this volume) note that "factual complementizers", which in our nomenclature correspond to semantically transparent propositional complementizers, are more borrowing-prone than "non-factual complementizers", which in our nomenclature mostly correspond to SoA-complementizers (e.g. the Balkan pro-infinitive markers). Put in the terminology and focus of the present volume, this means that complementizers with propositional scope are more susceptible to borrowing than complementizers with sub-propositional scope. We have too few complementizers of the SoA-type in the sample to generalize about its borrowability, but with the help of the figures from Table 6 we can submit a similar one-way implication concerning

semantic transparency: semantically neutral propositional complementizers are more prone to borrowing than semantically non-neutral complementizers encoding epistemic uncertainty.

6.4 Diachronic origins versus synchronic multifunctionality

The evidence for synchronic multifunctionality of complementizer forms (Section 5.1) is generally consistent with the diachronic information presented in this section. The only striking discrepancy is that although, as Table 3 indicated, semantically neutral complementizers are very often identical in form with adverbializers of purpose and reason, such adverbializers are missing from the list of their diachronic sources (cf. the discussion in Section 6.1). This leads us to the following conjectures: a) semantically neutral complementizers do not develop from adverbializers of purpose or reason; b) adverbializers of purpose and reason develop either from semantically neutral complementizers, or from the same sources as semantically neutral complementizers but indenpendently of them. In the latter case we are dealing with split paths in a polygrammaticalization cline (Craig 1991).

These conjectures are compatible with previous observations on development of complementizers; for instance, in the World Lexicon of Grammaticalization the path 'purpose' > 'complementizer' is lacking from the list of source–target pairs, but the reverse path 'complementizer' > 'purpose' is present in the list (Heine & Kuteva 2002: 319, 329). While complementizers give rise to purpose markers, purpose markers are often predecessors of cause/reason markers (e.g. Heine, Claudi & Hünnemeyer 1991: 157; van Gelderen 2010).

7 Summary

Despite the exploratory intent of the study, the analyses in the individual contributions allow us to make a number of tentative generalizations about the semantics of complementizers in the European languages.

The most pervasive contrast encoded in the complementation marking systems of the European languages is the contrast between propositions and state-of-affairs. The majority of languages convey this contrast within their complementation marking systems. In the rest of the cases we do not have evidence for this contrast, but in no case do we have evidence for the lack of the contrast between complement propositions and SoAs. The second most common semantic

contrast coded by complementizers concerns the distinction between uncertainty and epistemic neutrality. Based on the available evidence we claim that this distinction is found in the complementation marking systems of the majority of language families within the geographical boundaries of Europe.

Minor distinctions, coded in just a couple of languages of the sample, relate to the semantic domains of factuality of propositions, reality (or actuality) of states of affairs, evidentiality, quotativity, time, polarity, aspect and agent control, and to the domain of information structure.

In this conclusive chapter we discussed several structural phenomena with potential relevance to the semantics of complementizers. One such phenomenon was complementizer optionality, i.e. semantically neutral (or near-neutral) alternations between presence and absence of a complementizer in otherwise structurally identical complements. Complementizer optionality is a common phenomenon in the languages of Europe. Most of the European languages seem to conform to the following tendency: if a European language has an epistemically neutral complementizer and allows for complementizer omission, one of the omissible complementizers is the epistemically neutral one. The only exception is Irish, in which the non-neutral rather than the neutral complementizer is optional.

Also complementizer combination is quite common in the languages of Europe. Most languages seem to conform to the following tendency: if a language allows for complementizer combination and has an epistemically neutral complementizer, this complementizer occurs in complementizer combination(s). Again the only exception is Irish. Although complementizer combinations tend to be semantically insignificant, we have a couple of examples in the sample of semantically significant combinations.

The forms identified as complementizers tend to be multifunctional subordinators, but there are also items in the sample which function exclusively as complementizers and not as adverbializers or relativizers. The available evidence warrants the following statements. The epistemically neutral type of complementizers is more often formally identical with adverbializers and relativizers than is the uncertainty type of complementizers. Both the epistemically neutral type and the uncertainty type of complementizers are more often formally identical with adverbializers than with relativizers. The epistemically neutral type is most commonly identical with adverbializers of purpose, whereas the uncertainty type is most commonly identical with adverbializers of condition (protasis markers).

Forms functioning in dependent clauses as epistemically neutral complementizers may function in genuine main clauses as demonstrative pronouns, interrogative pronouns, proadverbs, particles, adpositions or coordinating conjunctions. Forms figuring in dependent clauses as complementizers encod-

ing epistemic uncertainty may also have main clause uses, but these are more uniform: they mostly function as polar question particles. In addition, all major semantic types of complementizers – epistemically neutral propositional complementizers, complementizers associated with SoAs, and uncertainty complementizers – occur in typical insubordination contexts.

The most common diachronic sources of semantically neutral propositional complementizers in the European languages are demonstrative, interrogative and relative pronouns, and proadverbs, whereas the uncertainty type of complementizers has as its sources polar question markers and conditional protasis markers. In relation to borrowability, epistemically neutral complementizers are more likely to be borrowed than complementizers conveying uncertainty.

With these general conclusions as well as with the more specific distributions identified in this conclusive chapter and in the individual contributions to the book, we hope to have paved the way for further studies of complementizer semantics.

References

Boye, Kasper. 2010. Reference and clausal perception-verb complements. *Linguistics* 48(2). 391–430.

Boye, Kasper. 2012. *Epistemic meaning. A crosslinguistic and functional-cognitive study* (Empirical Approaches to Language Typology 43). Berlin: Mouton de Gruyter.

Boye, Kasper & Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88(1). 1–44.

Boye, Kasper, Eva van Lier & Eva T. Brink. 2015. Epistemic complementizers: A crosslinguistic survey. *Language Sciences* 51(5). 1–17.

Bresnan, Joan W. 1970. On complementizers: Toward a syntactic theory of complement types. *Foundations of Language* 6(3). 297–321.

Craig, Collette G. 1991. Ways to go in Rama: A case study in polygrammaticalization. In Elisabeth Closs Traugott & Bernd Heine (eds.), *Approaches to Grammaticalization* (Typological Studies in Language 19[1]), 455–492. Amsterdam: John Benjamins.

Cristofaro, Sonia. 1998. Grammaticalization and clause linkage strategies: a typological approach with particular reference to Ancient Greek. In Anna Giacalone Ramat & Paul Hopper (eds.), *The limits of grammaticalization* (Typological Studies in Language 37), 59–88. Amsterdam & Philadelphia: John Benjamins.

Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.

Cristofaro, Sonia. 2007. Deconstructing categories: finiteness in a functional-typological perspective. In Irina Nikolaeva (ed.), *Finiteness. Theoretical and empirical foundations*, 91–114. Oxford: Oxford University Press.

Cristofaro, Sonia. 2014. Complementation hierarchies and the development of complement constructions. Paper presented at the workshop “Clausal complementation and (non)

factivity". 36. Jahrestagung der Deutschen Geselschaft für Sprachwissenschaft (DGfS), University of Marburg, 5–7 March.
Csató, Éva Á. 2010. Two types of complement clauses in Turkish. In Hendrik Boeschoten & Julian Rentzsch (ed.), *Turcology in Mainz/ Turkologie in Mainz* (Turcologica 82), 107–122. Wiesbaden: Harrassowitz.
Dahl, Östen. 1995. The marking of the episodic/generic distinction in tense-aspect systems. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The generic book*, 412–426. Chicago: University of Chicago Press.
Dik, Simon C. 1997. *The theory of Functional Grammar*. Part 2: *Complex and derived constructions* (Functional Grammar Series 21). Ed. by Kees Hengeveld. Berlin & New York: Mouton de Gruyter.
Dik, Simon C. & Kees Hengeveld. 1991. The hierarchical structure of the clause and the typology of perception-verb complements. *Linguistics* 29(2). 231–259.
Dixon, Robert M. W. 2006. Complement clauses and complementation strategies in typological perspective. In Robert M. W. Dixon, & Alexandra Y. Aikhenvald (eds.), *Complementation: A cross-linguistic typology* (Explorations in Linguistic Typology 3), 1–48. Oxford: Oxford University Press.
Donohue, Mark. 1999. *A grammar of Tukang Besi*. Berlin: Mouton de Gruyter.
Elsness, Johan. 1984. *That* or zero: A look at the choice of object clause connective in a corpus of American English. *English Studies* 65(6). 519–533.
Evans, Nickolas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness. Theoretical and empirical foundations*, 366–431. Oxford: Oxford University Press.
Frajzyngier, Zygmunt. 1995. A functional theory of complementizers. In Joan L. Bybee & Suzanne Fleischman (eds.), *Modality in Grammar and Discourse*, 473–502. Amsterdam: John Benjamins.
Giannakidou, Anastasia. 2001. Varieties of polarity items and the (non)veridicality hypothesis. In Jack Hoeksema, Hotze Rullmann, Víctor Sánchez-Valencia & Ton van der Wouden (eds.), *Perspectives on negation and polarity items* (Linguistics Today 40), 99–127. Amsterdam & Philadelphia: John Benjamins.
Givón, Talmy. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4. 333–377.
Givón, Talmy. 2001. *Syntax. An introduction*, vol. 2. Amsterdam & Philadelphia: John Benjamins.
Haspelmath, Martin. 1989. From purposive to infinitive – a universal path of grammaticalization. *Folia Linguistica Historica* 10. 287–310.
Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3). 663–687.
Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization. A conceptual framework*. Chicago & London: The University of Chicago Press.
Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
Heine, Bernd & Tania Kuteva. 2012. An integrative model of grammaticalization. In Björn Wiemer, Bernhard Wälchli & Björn Hansen (eds.), *Grammatical replication and borrowability in language contact* (Trends in Linguistics. Studies and Monographs 242), 159–198. Berlin: Mouton de Gruyter.
Hengeveld, Kees & J. Lachlan Mackenzie. 2008. *Functional Discourse Grammar. A typologically based theory of language structure*. Oxford: Oxford University Press.

Kaltenböck, Gunther. 2009. Is *that* a filler? On complementizer use in spoken object clauses. *Vienna English Working Paper* 18(1). 28–63.
King, Gareth. 1993. *Modern Welsh. A comprehensive grammar*. London & New York: Routledge.
Korhonen, Riitta. 1993. *Buts about conjunctions. A syntactic study of conjunction expressions in Finnish*. Helsinki: Suomalaisen Kirjallisuuden Seura.
Legate, Julie Anne. 2010. On how *how* is used instead of *that*. *Natural Language and Linguistic Theory* 28. 121–134.
Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
Menz, Astrid. 1999. *Gagausische Syntax. Eine Studie zum kontaktinduzierten Sprachwandel* (Turcologica 41). Wiesbaden: Harrassowitz.
Nordström, Jackie 2010. *Modality and subordinators*. Amsterdam: John Benjamins.
Roussou, Anna. 2010. Selecting complementizers. *Lingua* 120(3). 582–603.
Salminen, Tapani. 2002. Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. In Tatjana Agranat & Olga Kazakevič (eds), *Lingvističeskij bespredel: sbornik statejk 70-letiyu A. I. Kuznecovoj* [Linguistics without limits. Felicitation volume for A. I. Kuznecova on the occasion of her 70th birthday], 44–55. Moscow: Izdatel'stvo Moskovskogo universiteta.
Sands, Kristina. 2000. *Complement clauses and grammatical relations in Finnish*. Canberra: Australian National University PhD thesis.
Shank, Christopher & Hubert Cuyckens. 2010. Moving beyond synchrony: Applying a diachronic corpus-based multivariate analysis to examine the development and use of *I + think* as an 'epistemic parenthetical'. Paper presented at Societas Linguistica Europaea 2010, Vilnius, 1–5 September.
Sørensen, Marie-Louise L. & Kasper Boye. 2015. Vidensprædikatskomplementering [Knowledge-predicate complementation]. *Ny Forskning i Grammatik* 22.
Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.
van der Wouden, Ton. 1997. *Negative contexts. Collocation, polarity and multiple negation*. London & New York: Routledge.
van Gelderen, Elly. 2010. Features in reanalysis and grammaticalization. In Elisabeth Closs Traugott & Graeme Trousdale (eds.), *Gradience, gradualness and grammaticalization* (Typological Studies in Language 90), 129–148. Amsterdam & Philadelphia: John Benjamins.
Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca & New York: Cornell University Press.
Ylikoski, Jussi. 2003. Defining non-finites: action nominals, converbs and infinitives. *SKY Journal of Linguistics* 16. 185–237.

Language index

The language index adheres, where possible, to the language names and genetic classifications found in Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the World, Eighteenth edition*. Dallas, Texas: SIL International. Online version. http://www.ethnologue.com. Names of individual languages are given in regular; names of language families and subfamilies in italics.

Adyghe (*North Caucasian*) 6, 321, 691, 692–95, 697–699, 702–704, 707–711, 713, 718–721, 723, 725–732, 734, 736–739, 816, 827, 831–832, 840–841, 844, 849, 861, 871
Afrikaans (*Indo-European*) 6, 132, 144–146, 851, 870
Afro-Asiatic 6
Agul (*North Caucasian*) 746–747
Akhvakh (*North Caucasian*) 746
Albanian (*Indo-European*) 265–271, 273–277, 279–286, 288–290
– Gheg Albanian (Geg Albanian) 266, 277–279
– Tosk Albanian 266, 277–279, 281
Altaic 295, 530, 532, 567, 665–666
Ancient Greek (*Indo-European)* 287, 575
– Classical Greek 248
Andi (*North Caucasian*) 746
Andic (*North Caucasian*) 746
Angolar (*Creole*) 76
Aquitanian 379–380
Arabic (Afro-Asiatic) 290, 370, 413–414, 421–422, 424, 440
Arabic, Standard (*Afro-Asiatic*) 414, 421
Archi (*North Caucasian*) 6, 746, 778, 781, 786
Aromanian (*Indo-European*) 6, 265–267, 282, 287–289
Ashti (*North Caucasian*) 746
Avar (*North Caucasian*) 6, 746, 750, 754, 758, 768, 774, 780–782
Avar-Andic (*North Caucasian*) 6, 746
Awa-Cuaiquer (Awa Pit) (*Cuaiquer, Barbacoan*) 2
Azerbaijani (Azerbaijanian) (*Turkic*) 619, 622
Bagvalal (*North Caucasian*) 6, 746, 778–780, 782, 784, 785, 787–788, 844, 849
Balkan Romance (*Indo-European*) 265, 266, 270, 341
Balkan Slavic (*Indo-European*) 265–266, 271, 341, 818, 872
Balkan Turkish (*Turkic*) 287
Baltic (*Indo-European*) 6, 225–227, 238, 245, 251, 253–255, 262, 457–458, 490, 820, 823, 829, 838, 851, 859
Bantu (*Niger-Congo*) 93
Basque (*Isolate*) 6, 379–386, 389–392, 394, 395–399, 401–402, 404–407
Bats (Batsbi, Tsova-Tush) (*North Caucasian*) 6, 746
Bezhta (*North Caucasian*) 746
Bothlik (*North Caucasian*) 746
Breton (*Indo-European*) 20
Budukh (*North Caucasian*) 746
Bulgarian (*Indo-European*) 175, 178, 198, 211–220

Cafundo Creole (*Creole*) 76
Catalan (*Indo-European)* 76, 120, 723
Celtiberian (*Indo-European*) 20
Celtic (*Indo-European*) 6, 15–18, 20–22, 25, 44, 51–52, 62, 72, 815, 818, 820–821, 823, 839, 842, 851, 854–855
Central Romance (*Indo-European*) 77, 119, 121, 123
Chamalal (*North Caucasian*) 746
Chantyal (*Sino-Tibetan*) 250
Chavacano (*Creole*) 76
Chechen (*North Caucasian*) 6, 746, 778, 780–783
Chirag (*North Caucasian*) 746
Continental Celtic (*Indo-European*) 20

Cornish (*Indo-European*) 20
Croatian 178; see also *Serbian/Croatian*

Daghestanian (*North Caucasian*) 746
Danish (*Indo-European*) 6, 131–132, 134–135, 140–141, 143–152, 154–156, 159–162, 164–165, 167–169, 793, 798, 800, 802–804, 806, 813, 816, 820, 851, 854, 857, 861, 863, 870–871
Dargwa (Akusha) (*North Caucasian*) 6, 746
– Itsari Dargwa (Icari Dargwa) 746, 777–778, 780, 787–789
– Xuduc Dargwa 316
– Qunqi Dargwa 316, 746
Dido (Tsez) (*North Caucasian*) 6, 746, 754, 768–769, 781, 787
Domari (*Indo-European*) 370
Dutch (*Indo-European*) 6, 132, 134–135, 139, 143–144, 150, 152–154, 156–157, 163, 171, 203, 851, 863, 870–871

East Caucasian (*Nakh-Daghestanian*) (*North Caucasian*) 6, 297, 745–747, 754, 773–774, 777–778, 780–784, 786–787, 789–790, 830, 844, 859, 871, 873
East Slavic (*Indo-European*) 178
Emilian (*Indo-European*) 89
English (*Indo-European*) 1, 4, 6, 7–9, 18, 24, 27, 30, 34, 46, 50–51, 54, 57, 59, 75, 78, 85, 106, 123, 131–135, 137–138, 143, 144–149, 152–154, 156, 159–160, 162–169, 171, 203, 232, 258, 265, 272, 274–275, 295, 297–298, 300, 302, 318, 329, 396, 413, 427, 450–451, 453, 455–456, 461, 479, 481, 493, 499, 518, 542, 547, 589, 594, 619–620, 625–626, 638, 658, 675, 697, 703, 711, 722, 725, 726–727, 748, 752–753, 768, 795, 796, 800, 803, 811, 813, 818–820, 822–824, 833–839, 843, 847, 851–852, 854, 861–866, 870–871
– American English 168
Estonian (*Uralic*) 449–450, 452–460, 462–463, 468, 470, 471, 472–478, 480–482, 484–488, 490, 493–494

Fa D'ambu (*Creole*) 76
Faroese (*Indo-European*) 6, 132, 135, 139, 144, 156, 171, 795–796, 851, 863, 870, 871
Finnic (*Uralic*) 6, 449–454, 456, 458, 460–463, 466, 468, 470, 472, 474–477, 479, 483–484, 486, 489–494, 523–525, 815, 818, 820, 823, 829, 834, 846, 851–853, 859, 863, 873
Finnish (*Uralic)* 357, 371, 449, 450, 452, 454–465, 468–471, 473–474, 476–477, 479–483, 485, 488, 490, 492–493
Finno-Ugric (*Uralic*) 297, 449–500, 530–531, 587, 709, 833, 859
French (*Indo-European*) 6, 36, 75–81, 84–88, 90–91, 95–98, 100–107, 109–110, 114–116, 118–124, 285, 380, 684, 861, 863–864, 866–868, 870–871
Frisian (*Indo-European*) 6, 132, 144, 851, 870

Gagauz (*Turkic*) 622, 872
Galician (*Indo-European*) 76
Gapshima-Butri (*North Caucasian*) 746
Gaulish (*Indo-European*) 20
German (*Indo-European*) 6, 31, 77, 131–132, 134, 135, 141–144, 153–154, 157, 159, 163–164, 168, 225, 226, 253, 265, 322, 348, 427, 490, 493, 518, 522, 587, 622, 638, 796, 802, 820, 850–851, 863, 870, 872
Germanic (Indo-European) 6, 15, 25, 52–53, 56, 72, 95, 116, 120, 124, 131–136, 142, 144–145, 149, 152, 154–156, 159, 161–162, 169–171, 204, 289, 384, 449, 452, 490, 518, 522, 593–594, 815, 818, 820, 823, 848, 850, 852–854, 859, 862
Ghodoberi (Godoberi) (*North Caucasian*) 6, 746, 754, 780–781
Gothic (Goth.) (*Indo-European*) 134–135
Greek (*Indo-European*) 6, 265–269, 271–275, 279, 283–289, 341, 357, 362, 368, 370, 373, 826, 849, 863, 868, 870, 872
– Byzantine Greek 287
Greenlandic Inuktitut (*Eskimo-Aleut*)
– West Greenlandic 798
Guadeloupean Creole French (Guadeloupean) (*Creole*) 76

Haitian (*Creole*) 76
Hidatsa (*Siouan-Catawban*) 795, 798
High German (*Indo-European*) 449
Hinukh (Hinuq) (*North Caucasian*) 6, 745–751, 753–754, 758, 761, 763, 765, 767–770, 773–774, 776–777, 779–781, 783–787, 789–790, 815, 817–818, 823, 829–831, 837, 849–851, 853, 860, 870
Hungarian (*Uralic*) 6, 354, 357–358, 370, 560, 587–595, 597, 601–602, 606–607, 609, 612–615, 810, 815, 818, 851, 853, 859, 861, 863, 866, 872

Icelandic (*Indo-European*) 6, 132, 134–138, 140, 144, 150–154, 156, 159, 160, 166, 171, 474, 851, 863, 870–871
Indo-European 3, 6, 13, 20–21, 75, 225, 253–254, 265, 289–290, 293, 302, 490, 492, 523, 530, 619–620, 625–626, 859
Indo-Iranian (*Indo-European*) 6, 302
Indo-Portuguese (*Creole*) 76
Ingrian (*Uralic*) 479
Ingush (*North Caucasian*) 6, 746, 780–783, 787
Insular Celtic (*Indo-European*) 20
Irish (*Indo-European*) 6, 15–23, 25, 27–30, 32–35, 38, 41, 43–46, 48–53, 56–57, 59–69, 71–72, 815–816, 819, 835, 839, 842, 847, 849, 851, 853–856, 861, 863–864, 866–867, 871, 875
Italian (*Indo-European*) 6, 76, 78, 81, 84–89, 91–92, 95–97, 100–101, 104–108, 110–113, 115, 119–122, 178, 351, 357, 360, 362, 371, 413, 427, 861, 863–864, 869–872

Jakelteko (Jacaltec) (*Mayan*) 1, 93, 238, 802–803
Japanese (*Japonic*) 92–93, 158, 295, 802

Kabardian (*North Caucasian*) 703, 714, 719
Kabuverdianu (*Creole*) 76
Kadar (*North Caucasian*) 746
Kajtag (*North Caucasian*) 746
Kalmyk-Oirat (Kalmyk) (*Mongolic*) 6, 665–667, 671–673, 681, 709, 810, 813, 818, 832–833, 849, 871
Karakalpak (*Turkic*) 622
Karata (*North Caucasian*) 746
Karelian (*Uralic*) 6, 449, 452, 454, 457, 460, 463, 471, 473, 476, 482, 488
– Northern Karelian (North-Karelian) 449–450, 452, 455–457, 459, 470, 472, 482, 487, 502, 861, 863, 870
– Livvi-Karelian (Olonets Karelian) 449–450, 459, 482, 861–863, 870, 872
Karipuna Creole French (Karipúna Creole French) (*Creole*) 76
Kazakh (*Turkic*) 622
Khanty (*Uralic*) 587
Khinalug (*North Caucasian*) 746, 781
Khorasani Turkic (Khorasan Turkic) (*Turkic*) 622
Khvarshi (Khwarshi) (*North Caucasian*) 6, 746, 761, 778–781, 785
Kinya Rwanda (*Niger-Congo*) 93–94
Kipchak (*Turkic*) 529, 621
Komi (Uralic) 529–531, 533, 535–539, 542–543, 545–554, 557–581
Komi-Permyak (*Uralic*) 6, 529–531, 533–534, 537, 539, 542–543, 547–553, 559, 562, 577, 580
Komi-Zyrian (*Uralic*) 529–531, 533–539, 542- 548, 550–551, 553–559, 561–567, 569–572, 575–579, 580–581
Korean (*Koreanic*) 801
Korlai Creole Portuguese (*Creole*) 76
Kryts (Kryz) (*North Caucasian*) 746, 781, 872
Kubachi (*North Caucasian*) 746
Kurdish (Indo-European) 370
Kyrgyz (Kirghiz) (*Turkic*) 622

Lak (*North Caucasian*) 746–747
Lango (*Nilo-Saharan*) 2, 157–158
Late Latin (*Indo-European*) 122
Latgalian (*Indo-European*) 225, 254
Latin (*Indo-European)* 75, 120, 122–124, 379, 502–504, 548
Latvian (*Indo-European*) 3, 6, 225–237, 240–241, 243–246, 251, 253–261, 372, 457, 490, 813–815, 817–820, 836, 848–849, 853, 861, 863, 866–868, 870–872
Lega-Shabunda (*Niger-Congo*) 158

Lepontic (*Indo-European*) 20
Lezgi (Lezgian) (*North Caucasian*) 746–747, 780–785, 788
Liv (Livonian) (*Uralic*) 226, 449
Louisiana Creole (Louisiana Creole French) (*Creole*) 76, 88
Low German (*Indo European*) 449
Ludian (Lude) (*Uralic*) 6, 449, 453, 455–457, 459, 470, 483, 863, 871–872
Luxembourgish (Frankish) (*Indo-European*) 12

Macanese (*Creole*) 76
Macedonian (*Indo-European*) 6, 178, 265–268, 271–272, 276, 287–289
Malaccan Creole Portuguese (*Creole*) 76
Maltese (*Afro-Asiatic*) 6, 413–414, 416–419, 424, 427, 440, 445, 813, 815, 818, 820–821, 823, 837–838, 851–853, 863–865, 871
Mansi (*Uralic*) 545, 587, 614
Manx (*Indo-European*) 20
Mari (*Uralic*) 532, 539–540, 563, 576
Mayan 1, 93
Megleno-Romanian (*Indo-European*) 265
Middle English (*Indo-European*) 160
Middle French (*Indo-European*) 119
Mjurego-Gubden (*North Caucasian*) 746
Mongolic 6, 539, 666, 709
Mordvin (*Uralic*) 532
Morisyen (Creole) 76
Mugi (*North Caucasian*) 746
Muiri (*North Caucasian*) 746

Nakh (*North Caucasian*) 6, 746, 782
Neo-Aramaic (*Afro-Asiatic)* 370
Nogai (Noghay) (*Turkic*) 619, 621–625, 627–631, 635, 637–641, 646–658
Norwegian (*Indo-European*) 6, 132, 134–135, 143–144, 148, 150–151, 154, 156, 169, 450, 500, 502–503, 851, 863, 871

Old Chuvash (*Turkic*) 529
Old Danish (O.Da.) (*Indo-European*) 135
Old English (OE) (*Indo-European*) 134–135, 155, 160
Old French (*Indo-European*) 119
Old High German (OHG) (*Indo-European*) 134–135
Old Icelandic (*Indo-European*) 136
Old Indo-Aryan (Indo-European) 370
Old Lithuanian (*Indo-European*) 225, 251–254, 870
Old Norse (ON) (*Indo-European*) 134–135
Old Norwegian (O.Norw.) *(Indo-European*) 135
Old Romanian (*Indo-European*) 122, 289
Old Saxon (*Indo-European*) 135
Old Sicilian (*Indo-European*) 413
Old Swedish (O.Sw.) (*Indo-European*) 135
Old Zyrian *(Uralic*) 529, 531, 545, 551, 581, 861
Ossetic (*Indo-European*) 6, 293, 295–297, 302–306, 309, 317–318, 324–325, 327–329, 332, 334, 691, 738–739, 813, 815, 817–818, 827–828, 832, 834, 841–842, 845, 849, 858–859, 863, 866, 871

Palenquero (*Creole*) 76
P-Celtic (*Indo-European*) 20
Permian (*Permic*) *(Uralic)* 6, 529, 813, 815, 817–818, 822, 824, 829, 851, 853, 856, 859, 873
Persian, Iranian (*Indo-European*) 573, 619, 641, 781, 872
Polish (*Indo-European*) 6, 94, 136, 175, 177–178, 198, 205–211, 213, 215, 219–220, 225–226, 257, 259, 347–348, 351, 355, 359, 372, 813, 815, 818, 822, 846, 849, 853, 859–860, 863, 870–872
Portuguese (*Indo-European*) 6, 75–76, 78, 84–88, 95–96, 99, 100, 104–109, 115, 120–121, 861, 863–864, 870–871
Principense (*Creole*) 76
Proto-Finno-Ugric (*Uralic*) 490
Proto-Germanic (*Indo-European*) 134
Proto-Permian (Proto-Permic) (*Uralic*) 529

Q-Celtic (*Indo-European*) 20

Réunion Creole French (*Creole*) 76
Romance (*Indo-European*) 6, 32, 52, 62, 72, 75–80, 83–86, 88–89, 91–92, 95–96,

100, 102–103, 105, 107–111, 114–116, 118–124, 289, 379, 400, 402–403, 407, 413–414, 452, 815, 818, 820, 823, 848, 851, 859, 864, 869
Romani (*Indo-European*) 6, 265–266, 286–288, 341–362, 365–367, 369–374
– Arli, Serbia, Macedonia 347, 353, 363, 370
– Ayia Varvara, Greece 287
– Balkan 288
– Bergitka Romani, Poland 345, 347, 359, 368, 372, 846
– Crimean, Ukraine, Russia 350, 353–355, 358, 362
– Čuxny, Latvia, Estonia 345, 362, 365, 367
– Czech Vlax, Czech Republic 346, 354
– East Finnish Romani, Finland 346, 348, 366, 371
– East Slovak, Slovakia 348, 351, 356–357, 366, 371, 857
– Gimpeny, Ukraine
– Goli Cigani, Bulgaria 289
– Gurbet, Croatia, Serbia, Hungary 346, 350, 352, 353, 355, 363, 369
– Gurvari, Hungary 353–354, 358, 369, 371
– Kalajdži, Bulgaria 351, 368, 370
– Kalburdzhu, Bulgaria 371
– Kaldaraš, Romania 370
– Kalderaš, Romania 360
– Keremidarea, Romania 346
– Kishinjovcy, Ukraine 346
– Kosovan Romani, Serbia 369
– Kovacki, Macedonia 350
– Kubanski Servi, Ukraine 349, 365
– Kurturare, Romania 365
– Laeši Kurteja, Moldova 353, 356
– Lithuanian Romani, Lithuania 363–364, 368
– Lombard Sinti, Italy 357, 360, 371, 872
– Lotfitka, Latvia 366–367, 372, 872
– Lovari, Hungary 353–355, 358, 362
– Manuša Čurjaja, Croatia 355
– Mečkari, Albania 351, 353, 358–359
– Mexican Vlax, Mexico 353, 364
– Molise Romani, Italy 369
– Muzikantska Roma, Bulgaria 363
– Piculesti, Romania 365
– Plasčuny, Ukraine 356, 359
– Polish Xaladytka, Poland 348, 351, 355, 359, 372, 846
– Polska Roma, Poland 348, 359, 372
– Rakarengo, Romania 365
– Romacilikanes, Greece 364, 349, 352–353, 356
– Romungro, Slovakia, Hungary, Romania 348, 356–357, 359,
– Russka Roma, Russia 366–367, 371
– Servi, Ukraine 349, 353, 365, 371–373, 822
– Sinti, Romania 347–348, 351–352
– Sofades, Greece 352, 357, 363, 368, 371
– Sofia Erli, Bulgaria 349, 363
– Ursari, Romania 353, 359, 369, 370–371
– Vlax Romani, Bosnia, Romania, Serbia, Albania, Hungary 287
– Xandžari, Ukraine 349, 355, 358,
– Xoraxani, Bulgaria 351, 354, 362, 364, 366, 372, 872
Romanian (Daco-Romanian, Daco-Rumanian) (*Indo-European*) 6, 75–78, 84–89, 95–96, 100, 103–109, 116–124, 203–204, 265–267, 270, 288–289, 365, 370, 813, 815, 817–818, 820, 854, 860, 863, 868–869, 872
Romansch (*Indo-European*) 76
Russian (*Indo-European*) 6, 94, 175–176, 178–179, 182, 184, 190, 196, 198–200, 202–220, 226, 259–260, 316, 329, 333, 347, 354, 357, 360, 365–366, 368, 371–372, 449, 453, 455, 457, 490, 493, 502, 505, 509, 518, 523–526, 529–532, 535, 536, 542, 544, 547–550, 552, 555, 558, 560, 562, 572–573, 575, 580–581, 638, 666, 719, 720–722, 725–727, 739, 748, 773, 815, 817–818, 822, 827–829, 836, 845–849, 853, 856, 861, 863, 866, 870–872
Rutul (*North Caucasian*) 746

Saami, Akkala (*Uralic*) 501, 526
Saami, Inari (*Uralic*) 501, 525–526
Saami, Kildin (*Uralic*) 6, 499–506, 509, 511–527, 810, 824, 861, 863–864, 872
Saami, Lule (*Uralic*) 501, 524
Saami, North (*Uralic*) 6, 499–504, 510, 512–527, 861, 863–864, 872

Saami, Pite (*Uralic*) 501, 524
Saami, Skolt (*Uralic*) 6, 499–504, 507–509, 511–527, 861, 863–864, 872
Saami, South (*Uralic)* 501, 524
Saami, Ter (*Uralic*) 501, 526
Saami, Ume (*Uralic*) 501
Saint Lucian Creole French (*Creole*) 76
Sami (*Saamic*) (*Uralic*) 6, 499–504, 512, 518, 521, 523–526, 820, 873
San Miguel Creole French (*Creole*) 76
Sãotomense (*Creole*) 76
Sardinian (*Indo-European*) 6, 84–86, 88–89, 95, 104–106, 115, 120–122, 863–864, 871
Scottish Gaelic (*Indo-European*) 20
Serbian (*Indo-European*) 178; see also *Serbian/Croatian*
Serbian/Croatian (*Indo-European*) 118, 203
Seselwa (*Creole*) 76
Slavic (*Slavonic*) (*Indo-European*) 6, 77, 94, 178, 206, 212–213, 357, 360, 523, 587, 711, 719, 737, 820, 823, 838, 851, 853, 864
Slovak (*Indo-European*) 178, 206, 859
South Slavic (*Indo-European*) 347, 357, 360, 868
Spanish (*Indo-European*) 6, 75–76, 78, 84–86, 88, 95–96, 98–101, 104–109, 115, 119–121, 285, 380, 391–392, 399, 609, 723, 861, 863–864, 870–871
Swedish (*Indo-European*) 6, 132–135, 141–144, 149, 152–154, 156–157, 162, 166–167, 225, 357, 371, 449, 476, 500, 503, 524, 851, 854–856, 863, 870–872

Tabasaran (*North Caucasian*) 746
Tatar (*Turkic*) 539, 568, 573
Tayo (*Creole*) 76
Tindi (*North Caucsian*) 746
Tsakhur (*North Caucasian*) 6, 746, 754, 777–779, 781–782, 784–787, 789, 859, 872
Tsezic (*North Caucasian*) 6, 746, 754, 761, 769
Tsudakhar (*North Caucasian*) 746
Tukang Besi (*Austronesian*) 3, 850
Tundra Nenets (*Uralic*) 614
Turkish (*Turkic*) 619–633, 635–658
Turkmen (*Turkic*) 622

Udmurt (*Uralic*) 6, 529–531, 533–542, 544, 546–549, 552–554, 556–561, 563–564, 566–574, 576–581, 819, 822–824, 852–853, 856, 859, 861, 863, 866, 871
– Besermyan Udmurt (Beserman Udmurt) 531, 544, 548, 566, 568, 573–574
Ugric (*Uralic*) 6
Upper Guinea Crioulo (Crioulo) (*Creole*) 76
Uralic 6, 295, 500, 523, 529, 532, 555, 587–588, 614
Uyghur (Modern Uyghur) (*Turkic*) 641
Uzbek (*Turkic*) 290, 622, 641

Veps (*Uralic*) 449, 492, 870–871
Vod (Votic) (*Uralic*) 449

Welsh (*Indo-European*) 6, 15–24, 26–28, 30, 32–34, 36, 39, 41–42, 44–46, 48–51, 53–54, 56–57, 59–60, 62, 68–72, 819, 820–822, 839, 842–843, 849, 851, 853–854, 861, 863–864, 866, 871, 878
West Caucasian (*North Caucasian*) 6
West Germanic (*Indo-European*) 144, 204
West Greenlandic see *Greenlandic Inuktitut*
West Slavic (*Indo-European*) 178, 206
Western Romance (*Indo-European*) 111, 114, 116, 124

Yakut (Sakha) (*Turkic*) 626
Yiddish (*Indo-European*) 6, 132, 134, 136, 144, 156, 863, 871–873

Subject index

ablative case 303–304, 331, 463, 534, 563–564, 631, 751
abstract suffix 748, 750–751, 753, 758–759, 776, 785, 790, 829, 850
accusative case 123, 131, 303, 491, 506, 534, 560, 563, 581, 620, 631, 666–667, 676, 816
achievement predicate 226, 427, 479, 538, 637, 645, 748, 759, 764
action noun 560, 536
activity 192, 194, 210, 434, 443–444, 462, 464, 469, 478, 482, 535, 538, 703, 717, 730–731, 752, 755, 767, 784–785
activity type 753, 753–754, 757, 769, 772–773, 778–779, 784–785, 790
adposition 122, 287, 390, 517, 625, 665, 866, 875
adpositional meaning 775
adverbial clause 7, 9, 16, 53–54, 56–58, 62–66, 70, 86, 132, 155, 182, 191, 197, 202, 245, 252–253, 256, 295, 304, 308–309, 315, 334, 369, 418, 420, 425–426, 428–430, 455, 470–471, 473, 506, 516–517, 522, 655, 666–667, 697, 708, 710, 737, 746–747, 773, 775–776, 804–805, 819, 860–863
adverbial clause marker 178, 219, 710
adverbial subordinator 4, 5, 53, 149, 152–155, 169, 192, 196, 218–220, 227, 240, 252, 256–257, 798, 801, 806, 861
adverbializer 8, 83, 88, 123, 158, 178, 455–458, 470–471, 476, 517, 654, 859–865, 874–875
agreement 206, 343, 348–349, 381–383, 398, 442, 530, 627, 666–667, 696, 699, 702, 704–705, 707, 710, 729–730, 746–747, 754, 756, 761, 777, 779, 832
allomorph 386–387, 446, 621, 667
anteriority 530, 653, 732–734, 775–776
argument structure 619, 626, 629, 631, 641, 656, 686, 707, 712, 760, 773, 203, 318, 328, 331, 491
aspect 4, 24, 77, 171, 220, 265, 271, 282, 296, 303, 331, 342, 349, 369, 424, 429, 433–434, 444–445, 469, 547, 625, 627, 635, 640–641, 649, 653–654, 657–658, 695, 698–699, 705, 707, 843, 849, 875
– marker 621, 628, 666, 699, 704, 706–707
– aspecto-temporal 635
– aspectual contrasts 220, 469, 487
– aspectual type 843
assertion 90, 133, 162, 168, 202, 220, 239, 245, 249, 273, 280, 293, 297–299, 313, 319, 322, 333, 681–682, 691, 701–702, 716, 736, 765, 772
assertive 90, 811, 112, 138, 163, 168, 230, 251, 602, 604, 606, 608, 621, 636–637, 679–682, 811
assertive utterance predicate 138
asymmetry assumption 90
asyndetic construction 524
attributive clause 655, 860
auxiliary inversion 385, 387–388

borrowing 15, 287–289, 345, 351, 357, 360–362, 366, 368, 370, 374, 476, 505, 507, 509, 523–526, 555, 573, 872–874
bound morpheme 619–621, 635, 638

canonical complementizer 75, 83–86, 88–89, 123, 136, 142, 149, 155, 197, 205, 266, 272, 274, 279–280, 285–286, 450, 452, 459, 492, 499, 505–507, 510–511, 524, 526, 750, 781–782, 790, 810, 824, 837, 845, 853, 858–859, 863, 865
case marker 77, 318, 327, 330–331, 342, 418, 493, 621, 625, 655, 665, 693–699, 706–712, 714, 719–721, 732, 736–738, 776
certainty 1, 4, 27, 93, 112, 137–139, 238–239, 315, 407, 480–481, 483, 488, 499, 523, 609, 615, 681, 769, 802–804, 811, 818, 820, 822, 850, 853, 862
clause status 178
clause union 695, 748, 752, 776
clitic 15, 77, 82–83, 94, 188, 206, 208, 210, 254, 259, 272, 327–328, 330, 342, 406, 414, 417, 442, 450–451, 454, 476, 500, 509, 513, 527, 536–537, 557–558, 581, 592, 594–596, 598–600, 602, 605–607,

609, 611–612, 615, 624–625, 632, 748, 750, 752–757, 759, 761–762, 764–766, 770–774, 776–777, 780, 782, 785, 788, 790, 810, 829–830, 856, 860, 863, 871
coercion 182–183, 458, 477, 799–801, 805–806
commentative predicate 236, 637, 644–645, 648, 678–679, 686, 748
communication verbs 405, 406
comparative construction 193, 240, 776
complement clause 15, 17, 32, 70, 83, 94, 117, 146–148, 175, 177, 179, 180, 182–183, 186–188, 194–196, 199, 203–204, 207–209, 214, 232, 235, 243, 250–251, 259, 272, 275, 296–305, 307, 309–310, 312–313, 315–326, 330–333, 345, 348, 350, 354–355, 373–374, 383, 385–386, 388–390, 392–394, 398–402, 407–408, 413, 419–420, 423–426, 429, 434, 440, 445, 450–456, 458–459, 462–463, 466, 470, 474, 477, 480, 487, 490, 492, 499, 505–506, 509–512, 514, 518–523, 526, 533, 536–537, 539–541, 546–547, 554–556, 558–559, 568, 578, 580–581, 588–589, 592, 594–612, 615, 619, 621, 623, 625–632, 635–639, 641, 643, 648–649, 651, 654, 657–658, 670–674, 681, 685, 691, 695, 698–700, 703, 707, 709–711, 715–720, 723–724, 728–730, 732–733, 736–737, 746, 748–766, 768, 771–778, 780, 782–784, 786–789, 802, 805, 824, 826–828, 833, 837–838, 841–842, 845–848, 856, 860, 862–863, 869
complementation marker see *complementizer*
complementation strategy xii, 227, 233, 253, 295–297, 306, 691, 698, 709, 714, 728, 738–741
complementizer
– combinability 8, 32, 88, 255, 357, 367, 470, 532, 539, 572, 824, 833
– complex 219, 288, 360, 367, 476, 485, 522, 590–593, 609–610, 820, 854
– definition of 83, 451
– epistemically marked 225, 238, 241, 244, 258, 262
– epistemically uncertain 15, 802
– evidential 243
– finite 15, 51, 66, 68, 133, 379, 381, 383, 385–386, 399, 404, 406, 407, 826, 838–839, 862
– irrealis 211, 213, 237, 254, 259–261, 457, 836, 848
– non-complementizing functions of 8, 16, 57, 62, 369, 654, 592, 613
– omission of 4, 204
– propositional 232, 235–236, 238, 246, 253, 260–261, 258, 315, 815, 817–818, 842, 846, 854, 857–858, 861–864, 869–871, 873–874, 876
– Q-complementizer 454, 459, 471–473, 475–476, 479, 480–481, 485–489
– quotative 248, 538, 556, 569, 570, 574–575, 581, 831–832, 870
– simplex 354, 357, 360, 374, 846, 857
– state-of-affairs 235–237, 253, 256, 258–259, 261, 261, 815, 817, 836, 846, 857
complement-taking predicate 24, 34, 38, 71, 133, 146, 164, 170, 225, 227, 238, 295–296, 335–336, 421–422, 427, 432, 440, 445, 512, 518, 523, 629, 674–675, 679, 691, 694–696, 698, 705, 709–711, 713, 715–716, 719–723, 725, 727–728, 730–731, 734–738, 790, 801, 806, 811, 824–825; see also *achievement predicate, commentative predicate, desiderative predicate, epistemic predicate, fearing, predicate of, knowledge predicate, impersonal predicate, modal predicate, objective modal predicate, perception predicate, phasal predicate, predicate of discussion, predicate of investigation, pretence predicate, propositional attitude predicate, utterance predicate, volitional predicate*
conditional
– clause 136, 429, 455, 509, 511, 516, 695, 863
– mood 17, 29, 32, 57, 187, 209, 452, 456–459, 471, 481–482, 488–489, 536, 559, 595, 604, 609, 829, 836

– particle 194, 347, 536, 539, 552, 558–559, 580–581, 824, 347, 536, 539, 552, 558–559, 580, 824
– word 453, 455, 457, 574
connectivity 622
constituent question see *wh-question*
content of knowledge 653
content syntax 797
control 29, 31, 82, 180–183, 220, 245, 276, 310, 329, 332–333, 344, 346, 352–354, 356–357, 360, 367, 374, 443, 485–486, 488, 622, 627–628, 663, 738, 749, 845–849, 857, 875
control construction 182, 627–628
converb 226–227, 316, 326, 530, 538–540, 559, 564, 567, 667, 671–672, 674, 676, 680, 682, 684, 693, 694, 711, 746–749, 751–753, 757–762, 765–766, 773–777, 780–781, 785, 790, 830, 832–833, 837, 860
converbial 619, 632, 635, 665, 667, 671, 680, 682
copular 628–631, 646, 648–651, 653, 655, 657–658, 840
coreferentiality 310, 311, 329–330, 333, 434, 441, 445, 466, 563, 738
correlate 160, 303–304, 310, 313–314, 317, 320–326, 333, 343, 357, 593, 783, 840, 846
correlative adverb 293, 317–318, 332–333, 317
correlative construction 211, 334, 739
correlative pronoun 175, 195, 199, 293, 295, 297, 305, 309, 313, 317–322, 329, 332–333, 828, 835, 841
counterfactual 61, 208, 312, 316, 320, 407, 695, 822
counterfactuality 103
CP 385, 398, 404, 620, 684–685

dative case 77, 135, 182, 303, 382, 540–541, 562, 564, 589, 627, 631, 669, 673, 747, 749, 759, 784, 798
declarative 21, 23, 69, 131, 162, 171, 178, 384, 387–388, 390, 392, 395–397, 399, 402, 404–405, 407, 439, 453, 509, 521–523, 526, 533, 535, 593, 595–596, 598, 768, 784, 787, 797–801, 811–812
default 26, 177, 186, 201, 210, 225, 231, 236, 260, 262, 329, 374, 397, 401, 403–406, 445, 483, 526, 683, 685–686, 688, 754
degrammaticalization 241
dental modal subordinator 266, 270–271, 273–286, 288–289
deontic modality 234, 253, 701, 868
deranked verb form 629, 834, 844, 853
deranking 7, 9, 651, 812, 824, 658
derivation 179, 253, 272–273, 342, 360, 621, 628, 641, 654–656, 761
desiderative predicate 25, 28, 45, 208, 229, 281, 402, 472, 676, 846–847, 862; see also *volitional predicate*
dialect 144, 244, 254, 266–267, 277, 289, 303, 341–360, 362–372, 374
dialectal variation 386
direct object 57, 201, 213, 295, 303, 328, 331, 383, 416–417, 436, 442, 505–508, 510, 515, 534, 552, 560, 563, 620, 631, 641, 692, 708, 755, 798
direct question see *question*
direct reference 642–643, 645
direct speech type 753, 755, 757, 763, 766, 784, 790
DMS marker see *dental modal subordinator*
DP-hypothesis 626
dubitative 106, 115, 133, 135, 142, 166, 169, 381, 390

embedded construction 632, 634; see also *embedding*
embedded question 102, 199, 512, 523, 527, 592, 598–599, 751, 782; see also *embedding, question*
embedding 276, 346, 369, 537, 546, 576, 588, 590, 601, 603, 619, 632, 636, 651, 666, 669, 773, 777
epistemic distancing 187, 194, 202–203, 220
epistemic modal 35, 95, 123, 161–162, 167, 171, 210, 234, 239, 480–481, 587, 607, 700, 764, 790, 811–812, 824, 835, 35, 95, 123, 161
epistemic neutrality 131, 133, 138, 142, 149, 170, 818–819, 820–824, 875

epistemic predicate 118, 137
epistemic uncertainty 25, 34, 72, 131, 169, 236, 239, 480, 803–804, 818, 835, 862, 865–866, 868, 871–874, 876
epistemicity 114, 219, 484, 595–596, 608; see also *epistemic distancing, epistemic modal, epistemic neutrality, epistemic predicate, epistemic uncertainty*
equi deletion 472, 777
equi predicate 428, 435, 440, 443–444
event 44, 65, 67, 70, 164, 177, 181, 183, 185, 191–192, 202, 207, 210, 213, 215, 220, 234, 273, 290, 293, 297, 300, 301–302, 307–312, 323–324, 329, 331–333, 344, 354, 356, 374, 427, 436, 581, 599, 637, 642–643, 647, 649, 651, 691, 697, 791, 703, 711, 716–717, 719–720, 724–725, 727–728, 731–732, 737–739, 753, 755, 763, 772, 775–776, 812, 834, 845, 847–848; see also *state-of-affairs*
evidential 148, 193–195, 239, 241–244, 246–249, 262, 343, 472, 487, 483, 554, 595, 600–601, 608, 612, 615, 651, 653, 658, 757, 811, 829, 831, 867–868
evidentiality 219, 239, 241–242, 478, 480, 482–484, 494–495, 530, 587, 595, 608, 612, 615, 651, 653–654, 658, 761, 802, 810, 828–829, 875
extraposition 43

fact 260, 293, 297–299, 301, 307, 333, 401, 587, 636–638, 642, 657, 691, 701, 714, 717, 722, 738–739, 766–767, 771, 782, 785, 800, 803, 812, 826, 828, 834; see also *fact type*
fact type 228, 532–535, 545, 547, 553–554, 561, 572, 576–577, 580–581, 753–754, 766, 769, 772–773, 779, 784–786, 790
factive 29, 111, 140, 176, 177, 192, 200, 227, 232, 236, 269, 274, 279, 295, 297, 299–300, 315, 319–321, 329, 390, 392, 395, 401, 403, 407, 409, 433–434, 439, 445, 467–468, 532, 602–603, 636, 642, 678–679, 686, 691, 694–695, 698, 705–707, 709, 713–724, 726–729, 734–741, 759, 784, 786, 804, 810–811, 816, 825, 827–828, 844
factual 25, 31–32, 44, 51, 100–101, 109, 111–112, 115, 160, 344–347, 350, 357–359, 361, 367–368, 370–374, 393, 487, 642, 805, 825–828, 835, 840–841, 849, 860, 873
factuality 32, 33, 95, 102–104, 106–109, 115–116, 123–124, 344–345, 351–352, 356, 360, 373, 487, 489, 825, 827–828, 835, 840–841, 849, 875
fearing, predicates of 121, 187, 208, 234–236, 434, 469, 681, 748, 761, 764, 836–838, 844
finite complementation 62, 266–268, 272, 275, 279, 307, 333, 385, 386, 450, 459, 462–463, 468, 470, 489, 491, 494, 589
finite predicate 515, 551, 634, 742, 756, 773
finiteness 281, 398, 483, 627, 663, 698, 700, 702, 704
focality 649, 658, 840, 849
functional etymology 620
functional nominalization thesis 626–627
function-assignment 793, 797–798, 801

general subordinator 123, 160
generative syntax theory 626
generic event 301, 307, 309, 329, 332–333, 697, 727, 738–739
genitive case 77, 298, 303, 305, 318, 328, 330–331, 360–361, 390, 450, 460, 491, 560, 563–564, 629–630, 747, 750, 772, 776, 778–779, 784, 788, 844
genitive subject 492
genitive-possessive construction 629
grammaticalization 119–120, 145–146, 148–149, 250–251, 349, 368, 370, 374, 475, 493, 540, 548, 555, 592, 614, 671, 683, 688, 711, 737–738, 757, 762, 766, 793–796, 806, 832, 852, 870, 874
– polygrammaticalization 874

habituality 424, 530, 650
harmonic combination 8, 51, 158, 161, 167, 261, 468, 818–819, 851, 855
HIAT conventions 622, 658
higher clause 658, 794
high-focal 649, 651

illocution 114, 167, 199, 657, 797, 799, 801
illocutionary 75, 90–92, 114, 161–167, 472, 474, 579, 798–799, 801–802
illocutionary force 22, 90–91, 131, 162, 178, 298, 398, 399, 472, 521, 587, 636, 832
illocutionary value 399
immediate perception 209, 215–216, 233, 302, 308, 478, 532, 534, 541, 547, 576, 537, 817; see also *object of immediate perception, perception predicate*
immediate perception predicate see *perception predicate*
impersonal predicate 350, 362, 364–367, 442, 486
indefinite 52, 66, 77, 179, 271, 517, 800
indefiniteness 52, 192, 199, 202, 219, 220
indicative 31, 66–67, 69, 75, 110–115, 118–121, 123–124, 135, 187, 209, 214, 220, 244–245, 266, 268–275, 278–282, 284–286, 288–290, 303, 312, 316, 324, 330, 349, 391–392, 408, 457, 482–483, 489, 511, 530, 595–598, 600, 602–605, 607–612, 615, 636, 669, 670, 672, 675, 680–682, 699, 767, 782, 787, 790, 816–817, 821, 871
indicative complement clause 268, 280, 286, 680, 682
indirect question 19, 25, 29, 46–47, 71, 131, 133, 168, 171, 178, 189–190, 193, 210, 215–218, 306, 315, 334, 346, 387, 396–397, 480, 484, 598, 695–696, 737, 780, 782, 853
indirect speech 18, 168, 247–248, 405, 484, 485, 488, 770, 830, 833, 849, 857
indirect speech act 736
indirective 553–554
infinitive 427
infinitive marker 175, 267
information structure 303, 333, 576–577, 588, 718, 728–729, 754, 824, 825, 840–842, 849, 857–858, 875
insubordination 283–285, 472, 701–702, 766, 801, 865–869, 876
interrogative clause 178, 521, 599; see also *question*
interrogative clitic 592, 594–596, 598–599, 600, 602, 605–607, 609, 611–612, 615
interrogative focus 632
intraterminality 649, 651
irrealis 32, 66, 94–95, 141, 162, 195, 211, 213–214, 220, 225, 229–230, 237, 240–241, 245, 253–254, 258–262, 290, 301, 315, 317, 319, 321, 332–333, 360, 368, 452, 457–458, 468, 536, 545, 551, 559, 597, 710–711, 718–719, 723–724, 728, 737, 739, 749, 764–765, 767, 784, 786–787, 821, 824, 836, 847–849

knowledge acquired 202, 220, 477, 480, 483, 488, 643
knowledge of a possible fact 637–638, 657
knowledge predicate 38, 138–139, 141, 143, 163–164, 208, 227, 344, 347, 426–427, 436, 439, 458, 467, 519, 541, 546, 560, 637–639, 644, 677–678, 811, 814, 819, 850, 878, 541, 771, 842, 862
knowledge representation 637

language contact 265, 282, 286–290, 341–342, 347, 357, 360–362, 365–366, 370–374, 413, 449, 490, 502–503, 530, 532, 544, 641, 870
locative case 269, 571, 614, 627, 631
locative expression 416, 534, 656–657
logophoric 781
long-distance agreement 746, 754, 756, 761

manipulation 232, 344–346, 352–354, 356–361, 373, 637–638, 847
manipulative 352, 358
– predicate 182, 208, 226, 231, 281, 330, 343, 436, 467–468, 472, 675, 680, 748–749, 811–813, 834, 842, 862
manner 188, 295, 297, 308, 331, 390, 428, 431–432, 613–615, 642–644, 675, 691, 696, 708, 710, 711, 718–719, 736–738, 762, 775, 816, 861, 862, 871
masdar 694, 700, 746, 748, 750, 752, 771–772, 776–779, 782, 784–786, 788–790, 844
matrix predicate 180, 204, 424, 429, 433, 436, 440, 463, 589, 595–603, 605–610, 612, 615, 620, 637–638, 701–702, 757, 761, 772, 785, 839

mental perception 296, 478, 637, 639, 648
mixed extended projection 626–627
modal predicate 8, 77, 141, 167, 203–204, 231, 234, 281, 291, 296, 335, 349–350, 352, 427, 440, 462, 637, 639, 645, 736, 748–749, 777, 784, 811–812
– deontic 230, 250
– complement clause 269, 284, 348, 427, 698
modal prospectivity 646, 657
modifying clause 429
mood
– mood choice 114, 220, 324, 594, 601–606, 681
– mood concord 452, 456, 458–459
morphologically bound subordinator 625, 654
morpho-phonemic variation 414, 418

negation 8, 22, 40, 69, 82, 142, 168, 180, 183–184, 187, 190, 197, 202, 213–214, 228, 234–235, 256, 259, 286, 298, 301, 302, 306, 328, 330–331, 385, 387–388, 390, 391, 398, 400, 407, 433, 475, 500, 520, 531, 548, 557, 567, 574–575, 598, 601, 603, 605, 608, 646, 676, 697, 699–700, 704–707, 716, 720–722, 800, 836–839
negative 15, 17–20, 23, 25, 30, 33–36, 40, 41–44, 47, 52–53, 55, 62–64, 66, 69, 100, 101, 135, 142, 165–166, 180, 183–184, 192, 197, 214, 239, 244, 247, 249, 256, 258, 298, 302, 351, 358, 362, 365, 367, 369, 388–389, 391, 399–400, 422–423, 426, 429, 433, 434, 437–440, 483, 512, 531, 538–539, 548, 557, 566–567, 595, 601, 603–608, 632, 725–727, 821, 837–839, 860, 866
negative polarity 52, 165–166, 400, 837, 849
nominal predicate 274, 646, 686
nominalization 79–80, 306, 329–331, 335–336, 462, 467, 469, 486, 530, 626, 627, 636, 702–705, 710, 800, 846
nominalizer 450, 462, 641, 667, 699–700, 706–707, 779–780
nominalizing affix see *nominalizer*
nominalizing strategy 625
nominative case 122, 131, 275, 303, 318, 328, 331, 342, 450, 560, 630–631, 667, 816
non-assertive 87, 123, 521–522, 679–682, 687
non-embedded construction 188, 634
non-factive 145, 177, 192, 222, 227, 229, 236, 261, 295, 298–299, 317, 319–321, 401, 433–434, 439, 444–445, 467, 636, 678–679, 686–687, 697, 714–715, 719, 723, 728, 749, 810–811, 825, 828
non-factual 31, 100–104, 106, 108, 110–111, 115, 123, 206, 344–345, 347, 349–351, 358, 367–369, 373–374, 825–828, 835, 840–841, 849, 853, 873
non-factuality 75, 102–104, 123, 351, 360
non-finite complementation 385–386, 450, 459, 462–463, 468, 470, 489, 491, 494, 589
non-finite construction 70, 451, 546, 621, 626, 635, 654
non-finite predicate 382, 450, 490, 494, 539, 559, 746, 748, 756, 760, 767–768, 777, 782, 785, 790
non-focal 650
non-indicative 312, 316, 457, 595–598, 600, 602, 605, 607–608, 611–612, 680–681, 755, 763, 767, 770, 773, 790
non-logophoric 463, 465, 846
non-specific subject 630

object of perception 186, 202, 220, 477–480, 483, 488, 643; see also *immediate perception, perception predicate*
object sentence 556, 563, 581
objective modal predicate 637, 639, 645
obligatory control 627
oblique object 201, 631
obviation 180, 182, 207, 220, 404–405
operator 396–398, 401, 404–407, 606, 607, 608, 795, 839
overt subject 275, 629

parataxis 64, 296, 305, 316, 321, 335–336, 474, 555, 697
parentheticals 475, 756
participial 242, 276, 331, 463–464, 468, 470, 480, 490–491, 532, 534, 537,

562–563, 576–579, 581, 627–628, 635–637, 646, 649–650, 653, 655, 657, 665, 667, 670, 677–678, 680, 774, 782, 788, 824, 834, 837, 840
participle 153, 241, 270, 304–306, 326, 330–332, 337, 450, 460, 463, 465, 478–480, 488, 491–492, 531, 534, 538, 549, 559, 562–567, 571, 576–577, 579, 581–582, 584, 623, 628, 636, 638, 651, 667, 669, 672, 674, 694, 705, 718, 748, 750–753, 759–760, 768–769, 771, 774–776, 779–780, 785, 790, 817, 832, 860
PCU-predicate 102; see also *knowledge predicate, perception predicate, utterance predicate*
perception predicate 43, 69, 138–139, 186, 208, 210, 215, 233, 226, 296, 300–302, 304, 329, 332, 335, 373, 385, 444, 467–468, 477–478, 480, 482–483, 488, 494, 497, 519, 541, 589, 602, 604, 607, 637–638, 643–645, 676, 718–719, 727, 735–738, 740, 751, 755, 759, 769, 780, 785, 811, 813, 815–817, 832; see also *immediate perception, object of immediate perception*
person 18, 21, 49, 110, 140, 145–146, 168, 180, 184–185, 191, 203, 206–207, 212, 220, 226, 239–240, 247–248, 250–251, 270, 289, 299, 305, 318, 342–343, 350, 363, 364–365, 382, 386, 431, 433, 442, 474, 484–485, 515, 530, 536, 548, 557, 559, 570, 576, 580, 596, 600, 656, 670, 681, 690, 699, 702–704, 734, 747, 765, 766, 786, 826, 832–833, 837, 871
pseudo-predicate 440, 442
phasal predicate 226, 281, 296, 304, 335, 344, 427, 432, 441, 538, 561, 564, 566, 637, 639, 645, 695, 698, 727–728, 736, 738, 751, 777, 780, 785
polar interrogative see *polar question*
polar question 16, 18, 22, 25, 29, 50, 51, 52, 60, 71, 141, 164–165, 167, 188, 199, 210, 216, 396–397, 426, 454, 476, 500, 509, 512–513, 521, 523, 537–538, 557, 592, 598–599, 787, 802, 818–819, 838; see also *question*
polar question marker 15, 18, 22, 30, 32–33, 46–47, 49, 51, 60, 71–72, 191, 217, 475, 819, 871, 876; see also *question particle*
polarity 4, 51, 166, 184, 379, 391–392, 399–401, 407, 601, 605, 607, 835, 837–838, 849, 875; see also *negative polarity*
possessive 21, 22, 44, 55, 327–328, 330–331, 500, 530, 534, 560, 563–564, 577, 588–589, 626–629, 650, 655, 666–667, 670, 676, 681, 696, 699, 703–707, 730, 824, 832
possessor 59, 303, 534, 563–564, 629, 749
possessum 563
postposition 81, 540, 554, 565, 759
postterminality 649, 653, 658
postverbial construction 628, 649, 657
postverbial marker 628
potential type 752, 766, 773, 790, 228
predicate see *achievement predicate, commentative predicate, complement-taking predicate, desiderative predicate, epistemic predicate, fearing, predicate of, knowledge predicate, impersonal predicate, modal predicate, objective modal predicate, perception predicate, phasal predicate, predicate of discussion, predicate of investigation, pretence predicate, propositional attitude predicate, utterance predicate, volitional predicate*
predicate of discussion 137
predicate of investigation 137
predicate of knowledge see *knowledge predicate*
predicate of perception see *perception predicate*
predicate of saying and telling see *utterance predicate*
predication 15, 79–81, 90–92, 123, 192, 219, 265, 273, 294, 345, 357, 368, 374, 450, 492, 533, 555, 578, 581, 619–620, 646, 674, 814, 834, 859
presumption 651, 654, 658
presupposition 293, 297–298, 300, 319, 320–322, 333, 401, 691, 716

pretence predicate 143, 244, 246, 433, 637, 645
progressive 24, 282, 261, 478, 649, 840
proposition 1–2, 4, 15, 17, 25, 27, 31–32, 35, 38, 43–44, 51, 72, 93–95, 112, 131, 133, 136–143, 145, 151, 154–155, 157, 159–160, 164, 166–167, 169–171, 186–187, 194, 225, 238–239, 246, 259–262, 266, 280, 281, 293, 297–302, 307–315, 317, 319–321, 325–326, 332–334, 338, 343–344, 351–352, 356–357,, 386, 393, 395, 379, 396, 402, 404–405, 407, 422–427, 433, 454, 463–464, 473, 477–478, 480, 483, 487, 490–491, 494, 499, 502, 519–521, 538, 556, 559, 580–581, 596, 599–600, 603–606, 608–609, 615, 619, 626, 636–637, 642, 662, 679, 686, 691, 703, 714–715, 717, 719–722, 725–728, 732, 737–739, 757, 759, 761, 763, 767–768, 771, 790, 797–799, 802–803, 810–822, 824–827, 829, 835, 848, 850, 866–868, 873–875
propositional attitude 16, 30, 34, 142, 261, 280, 483, 561, 605–609, 626
propositional attitude predicate 41, 69, 138, 184, 208, 226–227, 229, 259, 426, 433, 436, 439–440, 457, 443, 463–465, 467, 480, 491, 512, 518, 541, 637, 643, 647, 675, 686, 715, 722–723, 748–749, 763, 771, 780, 811, 816–817, 819, 822, 829–830, 833, 842, 853–854, 862
propositional complement 3, 151, 227, 229, 231, 232, 234–236, 239, 245, 251, 259, 261, 273, 281, 477–478, 574, 581, 637, 686, 737, 813–814, 816, 818
propositional complementizer see *complementizer*
propositional question 632, 634
prosodic features 147, 624
prospectivity 646, 657
protasis 29–30, 32–33, 52, 61, 88, 107, 136, 154, 156–159, 164–165, 168–169, 171, 300, 320, 455, 819, 847, 856, 862, 871–871, 875–876
purpose clause 64–66, 202, 252–253, 345, 351–352, 354, 356–360, 373–374, 428, 445, 671, 720, 766, 773, 781–782, 862
purposive 10, 53, 55, 58, 63, 65, 67, 86, 155, 157, 171, 304, 536, 540, 548–552, 561, 565, 567, 570, 573–574, 577, 581, 602, 627, 667, 748–749, 752, 757, 759–761, 765–767, 773–774, 776, 861, 863

question 47, 71, 133, 136, 138, 162–164, 168, 170, 189, 215–217, 306, 397, 399–400, 407, 454, 478, 598, 609, 678, 695, 763, 788; see also *embedded question, indirect question, polar question, propositinal question, wh-question*
question particle 46, 306, 523, 525, 537, 539, 557–558, 580–581, 866, 869; see also *polar question marker, wh-question marker*
question word 133, 135, 233, 269, 279, 421, 425, 445, 451, 454, 471, 506–507, 509, 511, 514, 522–523, 527, 534, 537, 539, 545, 547, 576, 833, 857.
quotative particle see *complementizer*

raising 400, 428, 459, 538, 568, 571, 752
– predicate 70, 428, 443–444
recteur faible 92
reference strategy 538, 569
register 19, 23, 34, 62, 206, 470, 494, 504, 529, 548–549, 588
relative clause 7, 9, 16, 32, 57, 59, 62, 64–65, 68, 72, 159–160, 171, 215, 232–233, 274, 295, 303–304, 342, 384, 389, 396–397, 411, 418, 420–421, 471, 517, 695, 708–709, 746–747, 774–777, 782, 860–861, 863–865
relativizer 4, 33, 48, 59, 86, 120, 122, 123, 146, 153, 159–160, 361, 370, 374, 389, 396, 718, 860, 865, 870
reported speech 403, 515, 541, 755, 757, 761–764, 766, 770–771, 780–781, 786, 780–782, 786, 790, 830

same subject construction 181, 207, 220, 350, 352, 355, 428, 783, 846
scope 66, 77, 82, 91, 103, 108, 111–112, 159, 161, 227–228, 276, 298–299, 382, 391, 399–400, 405, 464, 474, 476, 478–479, 495, 796–801, 824, 830, 833, 853, 861, 866–868, 873
semi-factive 25, 27, 51, 467, 602, 604, 606, 678
similative 456–458, 468, 471, 473, 481, 493, 535, 553–554, 575, 822, 829
SoA see *state-of-affairs*
speech act type 753, 786, 790, 797
sprachbund 265, 286
state-of-affairs 3, 43, 90, 131, 228–229, 231–232, 234–237, 239, 248, 251, 253, 256, 259–260-262, 281, 300, 386, 405, 477, 483, 489–490, 494, 580, 636, 811–818, 820–821, 824, 834, 836–837, 838, 840, 844, 846, 850, 857, 868, 873–874
static copular particle 631
stylistic contrasts 206, 220, 505
sub-illocutionary meaning 803–804
subject marking 42, 43, 629
subjunctive 57, 64–67, 75, 77, 80, 82, 110–115, 118–123, 166, 180, 188, 214, 219–220, 270–271, 273, 275, 279, 281–282, 285, 288–289, 303, 312, 314, 347, 349–350, 367, 374, 379, 386, 391–392, 399–407, 452, 456, 596, 598, 603, 609–610, 615, 777, 782–783, 787–788, 790, 822, 848
subjunctor 506, 522–526, 619, 637
subordinating conjunction 271, 421, 445, 450, 587
subordinator 4–5, 7, 11, 16, 53, 53–55, 62–64, 123, 128, 133, 149, 152–156, 160, 169, 171, 178, 191–192, 196, 200, 210, 218–220, 227, 236–241, 245, 252, 254, 256–258, 266, 272–275, 279, 285–287, 293, 295, 302–307, 306–325, 329, 332–333, 357, 389, 398, 408, 421, 452–453, 456–457, 459, 470, 486, 494, 496, 508, 522, 587, 592–594, 614–625, 641, 695, 727, 798–799, 801, 803, 805–806, 809, 834, 842, 845, 859–861, 875; see also *complementizer*
superordinate clause 175, 619, 631

tag question 91, 168, 788
tense 15–18, 22, 24–27, 32, 37–39, 41, 47, 49, 59, 77, 120, 168, 175, 182, 191, 203, 208, 232, 235, 250, 265, 268, 271, 275–276, 282, 299, 303, 328, 330–331, 342, 348–350, 373, 382, 386–387, 424, 429, 433–434, 441, 445, 460, 463–465, 475, 487, 530–531, 536, 543, 547–548, 550, 554, 557, 569, 588, 600, 625–627, 654, 666, 693–695, 698–700, 704, 706–707, 710, 732–735, 737–738, 756, 763–765, 770, 787, 824, 833–835, 837, 843.
time adverb 640, 645, 658
truth value 102, 131, 133, 161–162, 168–169, 266, 280, 297–301, 312, 345, 351, 356, 373, 386, 392–393, 406–407, 478, 499, 520, 580, 636, 724, 753, 768, 803, 814, 826

verb of volition see *volitional predicate*
verb sequence 427–428, 431, 475
verbal noun 17, 24, 26, 28–29, 37–38, 42–43, 46, 54, 56, 60, 69–70, 195, 327, 637–638, 651, 696, 727, 735–738, 844
verbal predicate 558
verbum dicendi 541, 619, 634; see also *utterance predicate, PCU-predicate*
verbum sentiendi 541; see also *perception predicate, PCU-predicate*
veridical 111, 266, 602–606, 608, 615
veridicality 601–609
volition 237, 402, 486, 494, 612, 836, 846, 867
volitional 229–231, 234, 254, 262, 441, 485–486, 488, 642, 845
volitional predicate 104, 180, 121, 229, 232, 351, 358, 402–403, 405, 598, 637, 639, 645, 648; see also *desiderative predicate*
– yes-no question see *polar question*

wh-element 79–80, 609, 632
wh-question 59, 199, 210, 306, 397, 527, 537, 635, 695, 768, 787, 789
– marker 18, 32, 454, 471, 514, 522
word order 9, 17, 21, 77, 164–165, 178–179, 195, 200, 218–219, 233, 306, 321, 342, 367–368, 373, 384–385, 388–389, 451, 458, 477, 509, 513, 523, 530, 540–541, 543, 546, 552, 556, 588, 624, 629, 666, 692, 747, 757, 798, 800, 810, 842
world evaluation 402

zero marking 748, 752–753, 755, 767–768, 770, 777, 781, 785