

The
Indo-European
Languages

Second Edition

Edited by Mate Kapović

ROUTLEDGE LANGUAGE FAMILY SERIES



THE

INDO-EUROPEAN

LANGUAGES

The Indo-European Languages presents a comprehensive survey of the individual languages and language subgroups within this language family.

With over four hundred languages and dialects and almost three billion native speakers, the Indo-European language family is the largest of the recognized language groups and includes most of the major current languages of Europe, the Iranian plateau and the Indian subcontinent.

Written by an international team of experts, this comprehensive, single-volume tome presents in-depth discussions of the historical development and specialized linguistic features of the Indo-European languages.

This unique resource remains the ideal reference for advanced undergraduate and postgraduate students of Indo-European linguistics and languages, but also for more experienced researchers looking for an up-to-date survey of separate Indo-European branches. It will be of interest to researchers and anyone with an interest in historical linguistics, linguistic anthropology and language development.

Mate Kapović is Associate Professor in the Department of Linguistics at the University of Zagreb, Croatia. He currently teaches courses on Indo-European phonology, Indo-European and Balto-Slavic accentuation, Indo-European morphology, and general phonology.

ROUTLEDGE LANGUAGE FAMILY SERIES

Each volume in this series contains an in-depth account of the members of some of the world's most important language families. Written by experts in each language, these accessible accounts provide detailed linguistic analysis and description. The contents are carefully structured to cover the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology, and sociolinguistics.

Every volume contains extensive bibliographies for each language, a detailed index and tables, and maps and examples from the languages to demonstrate the linguistic features being described. The consistent format allows comparative study, not only between the languages in each volume, but also across all the volumes in the series.

- | | |
|---|--|
| The Austronesian Languages of
Asia and Madagascar
<i>Edited by Nikolaus Himmelmann &
Sander Adelaar</i> | The Mongolic Languages
<i>Edited by Juha Janhunen</i> |
| The Bantu Languages
<i>Edited by Derek Nurse &
Gérard Philippson</i> | The Munda Languages
<i>Edited by Gregory D.S. Anderson</i> |
| The Celtic Languages 2nd Edition
<i>Edited by Martin J. Ball &
Nicole Müller</i> | The Oceanic Languages
<i>Edited by John Lynch, Malcolm Ross &
Terry Crowley</i> |
| The Dravidian Languages
<i>Edited by Sanford B. Steever</i> | The Romance Languages
<i>Edited by Martin Harris &
Nigel Vincent</i> |
| The Germanic Languages
<i>Edited by Ekkehard König &
Johan van der Auwera</i> | The Semitic Languages
<i>Edited by Robert Hetzron</i> |
| The Indo-Aryan Languages
<i>Edited by George Cardona &
Dhanesh K. Jain</i> | The Sino-Tibetan Languages
2nd Edition (forthcoming)
<i>Edited by Graham Thurgood &
Randy J. Lapolla</i> |
| The Indo-European Languages
2nd Edition
<i>Edited by Mate Kapović</i> | The Slavonic Languages
<i>Edited by Bernard Comrie &
Greville G. Corbett</i> |
| The Iranian Languages
<i>Edited by Gernot Windfuhr</i> | The Tai-Kadai Languages
<i>Edited by Anthony Diller</i> |
| The Languages of Japan and Korea
<i>Edited by Nicolas Tranter</i> | The Turkic Languages
<i>Edited by Éva Csató &
Lars Johanson</i> |
| The Khoesan Languages
<i>Edited by Rainer Vossen</i> | The Uralic Languages
<i>Edited by Daniel Abondolo</i> |

THE
INDO-EUROPEAN
LANGUAGES

Second edition

Edited by
Mate Kapović

 **Routledge**
Taylor & Francis Group
LONDON AND NEW YORK

Second edition published 2017
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2017 selection and editorial matter, editor; individual chapters, the contributors

The right of the editor to be identified as the author of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

This edition adapted from the 1998 Routledge translation of *Le Lingue Indoeuropee* © 1993 Società editrice Il Mulino, Bologna

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record for this book has been requested

ISBN: 978-0-415-73062-4 (hbk)

ISBN: 978-1-315-67855-9 (ebk)

Typeset in Times New Roman
by Apex CoVantage, LLC

CONTENTS

<i>List of illustrations</i>	vii
<i>Preface</i>	xiii
<i>Abbreviations</i>	xv
Indo-European languages – introduction <i>Mate Kapović</i>	1
1 Proto-Indo-European	
1 Proto-Indo-European phonology <i>Mate Kapović</i>	13
2 Proto-Indo-European morphology <i>Mate Kapović</i>	61
3 Proto-Indo-European syntax <i>Thomas Krisch</i>	111
2 Proto-Indo-European and language typology <i>Ranko Matasović</i>	153
3 Anatolian <i>H. Craig Melchert</i>	171
4 Indo-Iranian	
1 Indo-Iranian <i>Leonid Kulikov</i>	205
2 Indo-Aryan <i>Leonid Kulikov</i>	214
3 Iranian <i>Nicholas Sims-Williams</i>	263
5 Greek <i>Rupert Thompson</i>	287
6 Italic <i>Rex Wallace</i>	317
7 Celtic <i>Patrick Sims-Williams</i>	352
8 Germanic <i>Joshua Bousquette and Joseph Salmons</i>	387
9 Armenian <i>Birgit Anette Olsen</i>	421
10 Tocharian <i>Douglas Q. Adams</i>	452

11 Balto-Slavic

1 Balto-Slavic *Steven Young* 479

2 Baltic *Steven Young* 486

3 Slavic *Marc L. Greenberg* 519

12 Albanian *Alexander Rusakov* 552

Language Index 609

Subject-Author Index 617

ILLUSTRATIONS

MAPS

1.1	IE languages today	2
1.2	IE languages in the 1st millennium BCE	6
1.1	The <i>centum-satem</i> branches of IE	28
3.1	Anatolia in the Late Bronze Age	173
3.2	Anatolia in the Iron Age	174
5.1	The Greek dialects	288
6.1	The regions of ancient Italy	318
6.2	The languages of pre-Roman Italy	319
7.1	Latin provincial lapidary inscriptions bearing Celtic personal names	355
7.2	Density of Celtic-looking ancient place-names	356
9.1	Armenian in ancient and medieval times	423
10.1	Kucha (= Kuci) and its dependencies in Han times	453
11.1	Baltic tribes at the beginning of the second millenium ad	487
11.2	Map of Slavic languages	520
12.1	The Albanian dialects	553

FIGURES

1.2	Sentence scheme 1	116
1.3	Sentence scheme 2	116
1.4	Structure of Homer Iliad 1,8	117
1.5	Structure of Homer Odyssey 4,353	118
1.6	Structure of Homer, Odyssey 20, 48	122
1.7	Structure of Rigveda 1,32,2	133
1.8	Structure of Homer, Iliad 18,468 / 1	134
1.9	Structure of Homer, Iliad 18,468 / 2	135
1.10	Structure of Homer, Iliad 5,97	136
1.11	Structure of Homer, Iliad 5,99	136
1.12	Structure of Homer Iliad 5,100	137
1.13	Structure of Homer, Iliad 5,101	137
1.14	Hettrich 2002: 47 The spectrum of meaning of the Rgvedic instrumental	141
4.1	Chronology of Indo-Aryan languages and texts	215
4.2	Polyglossia in ancient India	216
4.3	Simple vowels	267
4.4	The first and second series of palatals in Iranian	270

5.1	Inherited vowel system of Proto-Attic-Ionic	294
5.2	Vowel system of Proto-Attic-Ionic after compensatory lengthening	295
5.3	Vowel system of classical Attic, showing official spellings after 403 BC	296
6.1	<i>Fibula Praenestina</i> (CIL 1 ² .3)	320
8.1	Two models of Germanic subgrouping	388

TABLES

1.1	PIE stops	14
1.2	The reflection of PIE palatovelars in <i>satem</i> languages	23
1.3	Basic reflexes of PIE stops	28
1.4	The reflexes of PIE asyllabic resonants	34
1.5	The reflexes of PIE syllabic resonants	36
1.6	The reflexes of PIE *CR _h HC	39
1.7	The reflexes of PIE *CR _h HV	39
1.8	The reflexes of PIE vowels	41
1.9	The reflexes of PIE diphthongs	51
1.10	IE <i>o</i> -stems (singular)	64
1.11	IE <i>o</i> -stems (plural)	64
1.12	IE <i>o</i> -stems (neuter)	65
1.13	IE <i>o</i> -stems (dual)	66
1.14	IE <i>eh</i> ₂ -stems (singular)	66
1.15	IE <i>eh</i> ₂ -stems (plural)	67
1.16	IE <i>eh</i> ₂ -stems (dual)	67
1.17	IE root nouns (singular)	71
1.18	IE root nouns (plural)	71
1.19	IE <i>i</i> -stems (singular)	72
1.20	IE <i>i</i> -stems (plural)	72
1.21	IE <i>u</i> -stems (singular)	73
1.22	IE <i>u</i> -stems (plural)	74
1.23	IE <i>r</i> -stems (singular)	75
1.24	IE <i>r</i> -stems (plural)	75
1.25	IE <i>n</i> -stems (singular)	76
1.26	IE <i>n</i> -stems (plural)	76
1.27	IE <i>s</i> -stems (singular)	78
1.28	IE <i>s</i> -stems (plural)	78
1.29	IE demonstrative pronoun *so – *seh ₂ – *tod (singular)	85
1.30	IE demonstrative pronoun *so – *seh ₂ – *tod (plural)	86
1.31	IE demonstrative pronoun *so – *seh ₂ – *tod (dual)	86
1.32	IE demonstrative pronoun *ey (*is) – *ih ₂ – *id (singular)	87
1.33	IE demonstrative pronoun *ey (*is) – *ih ₂ – *id (plural)	87
1.34	IE athematic present	93
1.35	IE thematic present	93
1.36	IE present dual endings	94
1.37	IE root-aorist	96
1.38	IE sigmatic aorist	96
1.39	IE thematic aorist	97

1.40	IE perfect	98
1.41	IE middle present	99
1.42	IE middle aorist	100
1.43	IE athematic subjunctive	101
1.44	IE thematic subjunctive	101
1.45	IE athematic optative	102
1.46	IE thematic optative	102
1.47	IE athematic imperative	103
1.48	IE thematic imperative	103
1.49	IE future imperative	103
3.1	Anatolian nominal endings	178
3.2	Anatolian personal pronouns	181
3.3	Anatolian anaphoric pronouns	182
3.4	Anatolian interrogative-relative pronoun	183
3.5	Hittite verbal inflection	185
4.1	Encoding transitivity oppositions: diachronic typological features of some language families	210
4.2	Vedic texts and schools	218
4.3	Abbreviations of texts (text sigla)	220
4.4	Old Indo-Aryan consonantism	222
4.5	Assimilation in consonant clusters before dental stops	230
4.6	The main alternation grades	231
4.7	Old Indo-Aryan case endings	233
4.8	Consonant declension	235
4.9	(<i>A</i>) <i>nt</i> -declension	235
4.10	<i>a</i> -declension (type <i>devá-</i>)	236
4.11	<i>ī</i> -declension (types <i>devī-</i> and <i>vṛkī-</i>)	237
4.12	PIE sources of case endings	237
4.13	Vedic verbal personal endings	239
4.14	Series of the PIE verbal endings	240
4.15	PIE sources of Vedic personal endings	240
4.16	The Vedic verbal system: examples	243
4.17	The Vedic verbal system: examples	244
4.18	The Vedic system of present stem types	245
4.19	The inventory of the present passive forms attested in the RV and AV	246
4.20	Passive paradigm in early Vedic	247
4.21	Declension of masculine <i>a</i> -stems (PIE stems in *o)	275
4.22	Declension of Av. <i>xratu-</i>	276
4.23	Secondary endings in Avestan	279
4.24	Primary endings in Avestan	280
4.25	Conjugation of the thematic present indicative active	280
4.26	Conjugation of the thematic present indicative middle	280
5.1	The Greek (Ionic) alphabet	289
5.2	Treatment of labiovelars in Attic-Ionic and West Greek	292
5.3	Forms of the first declension (<i>a</i> -stems) in classical Attic	298
5.4	Forms of the second declension (<i>o</i> -stems) in classical Attic	299

5.5	Underlying endings of third declension, and consonant stem forms in classical Attic	299
5.6	Forms of the third-declension <i>s</i> -stems, <i>i</i> -stems and <i>u</i> -stems in classical Attic	300
5.7	Underlying forms of the <i>-ēu</i> stems, and the resulting forms in classical Attic	301
5.8	Forms of the personal pronouns	304
5.9	Thematic and athematic primary and secondary endings of the active voice	306
5.10	Endings of the strong and weak aorist	307
5.11	Endings of the middle voice	308
5.12	Forms of the perfect and pluperfect in classical Attic	310
6.1	The phonemes of Classical Latin	325
6.2	Classical Latin consonant-stem endings	330
6.3	Classical Latin <i>o</i> -stem endings	330
6.4	Classical Latin noun paradigms, masculine and feminine genders	331
6.5	Classical Latin noun paradigms, neuter	331
6.6	Classical Latin ‘mixed’ declension	333
6.7	Classical Latin <i>ē</i> -stems	333
6.8	Latin personal pronouns	335
6.9	Classical Latin relative pronoun	336
6.10	Organization of verb system	336
6.11	Latin conjugation classes	337
6.12	Subjunctive mood	338
6.13	Personal endings	339
6.14	Perfect passives	339
7.1	Comparison of the ordinal numerals	353
7.2	The Common Celtic vowel system	362
7.3	Pre-Celtic stops	363
7.4	Common Celtic stops	363
7.5	Brythonic consonantal developments	365
7.6	Gaelic consonantal developments	367
7.7	Nominal stems in <i>-o-</i>	368
7.8	Nominal stems in <i>-ā-</i>	369
7.9	Nominal stem in <i>-n-</i>	369
7.10	A possible origin of Old Irish conjunct endings	373
7.11	A possible origin of Old Irish absolute endings	373
7.12	Old Irish passive forms (present tense)	374
7.13	Old Irish deponent forms (present tense)	374
8.1	Pre-Germanic consonants	391
8.2	IE and Germanic stops	392
8.3	Old High German weak verbs	396
8.4	Germanic strong verbs, classes I–V	398
8.5	<i>a</i> -stem masculine nouns	401
8.6	Examples of other noun classes	402
9.1	The Armenian alphabet	422
10.1	The development of PIE vowels in Tocharian	455
10.2	Tocharian consonants	458

10.3	Tocharian vowels	458
10.4	An example of Tocharian nominal declension	462
10.5	Some examples of Tocharian adjectival declension	464
10.6	Deictic pronouns in Tocharian	465
10.7	First and second person pronouns in Tocharian	466
10.8	Tocharian numbers from ‘one’ to ‘ten’	466
10.9	The declension of ‘one’ in the Tocharian languages	467
10.10	Conspectus of the Tocharian verb	469
11.1	The vowel system of common Baltic	489
11.2	The vowel system of East Baltic	490
11.3	The vowel system of Lithuanian	491
11.4	The vowel system of Latvian	491
11.5	Lithuanian accent paradigms for <i>a</i> -stem nouns	494
11.6	The Proto-Baltic consonant system	496
11.7	Declension of <i>a</i> -stem nouns	502
11.8	Declension of <i>a/ā</i> -stem and <i>u/i</i> -stem simple adjectives in Lithuanian	504
11.9	Declension of <i>a/ā</i> -stem and <i>u/i</i> -stem definite adjectives in Lithuanian	505
11.10	Declension of the Lithuanian third-person personal pronoun <i>jis, ji</i>	507
11.11	Proto-Indo-European to late common Slavic vowels	522
11.12	Late common Slavic vowels	523
11.13	PIE, Baltic, and Slavic accent paradigm correspondences	526
11.14	Accentual paradigms	527
11.15	Proto-Indo-European to late common Slavic consonants	528
11.16	Cluster simplification	530
11.17	First velar palatalization	530
11.18	Second and third velar palatalizations	531
11.19	‘Deiotation’ reflexes	532
11.20	Common Slavic consonant system	533
11.21	Vocalic stems	535
11.22	Consonantal stems	535
11.23	Vocalic stem endings	536
11.24	Consonantal stem endings	536
11.25	OCS 3 pers. pronoun	538
11.26	Interrogative pronouns	538
11.27	OCS adverbs of place, time, manner, direction	539
11.28	Inflection of non-past finite verb forms	540
11.29	Verb classes, following Leskien	540
11.30	Slavic verb form classes	541
11.31	Deverbatives from CSL. *vĕdĕti ‘to know’, derivatives of CSL. *ryba ‘fish’	543
12.1	Albanian alphabet	559
12.2	PIE short vowels (in stressed position)	563
12.3	PIE long vowels (in stressed position)	563
12.4	PIE diphthongs (in stressed position)	564
12.5	PIE sonorants	567
12.6	PIE obstruents	569
12.7	PIE *s, *y, *w	570
12.8	Noun declension in singular	574

12.9	Noun declension in plural	575
12.10	Articulated adjectives (unmarked order)	581
12.11	Articulated adjectives (inverted order)	581
12.12	Declension of the personal and demonstrative pronouns	582
12.13	Declension of the possessive pronouns	583
12.14	Cardinal numerals	584
12.15	Verb conjugation (the major conjugational classes)	589
12.16	Conjugation of the verbs <i>jam</i> 'to be' and <i>kam</i> 'to have'	591

PREFACE

This volume is formally the 2nd edition of Ramat and Ramat 1998, itself an English translation of Ramat and Ramat 1993. Although the idea and the scope of the volume have remained much the same, the volume itself has changed a lot (the new editor included), and not only due to the necessary updates, inescapable after almost a quarter of a century has passed since the first edition.

The chapter “Proto-Indo-European: Comparison and Reconstruction” (1st ed.) is now replaced by three completely new chapters (“Proto-Indo-European Phonology,” “Proto-Indo-European Morphology,” and “Proto-Indo-European Syntax”). Instead of the chapter “Sanskrit” (1st ed.), there is a new chapter “Indo-Aryan.” Also, a separate short prechapter “Indo-Iranian” has been added. Instead of the chapters “Latin” and “The Italic Languages” (1st ed.), there is only the new chapter “Italic” now. A new prechapter “Balto-Slavic” has been added to precede “Baltic” and “Slavic.” A new chapter “Fragmentarily Attested Indo-European Languages” was planned, but the text has unfortunately not been finished on time. The chapter “The Indo-Europeans: Origins and Culture” (1st ed.) is completely out. The order of the chapters is new. The only authors remaining from the 1st edition are Nicholas Sims-Williams (“Iranian”) and Patrick Sims-Williams (“Celtic”). Their chapters are updated versions of their chapters in the 1st edition; all the other chapters (written by different authors) are completely new. Some of the authors from the 1st edition have, unfortunately, passed away in the meantime, like the great Calvert Watkins (1933–2013) and Werner Winter (1923–2010). New maps and illustrations were also added.

The idea of the volume is to provide a compact but informative introduction to and overview of Indo-European languages, with the stress on Proto-Indo-European, the oldest well attested IE languages, and the largest IE branches. The aim of the volume is to provide an introduction to beginners or non-Indo-Europeanists, but also to be used as a reference book for scholars already doing research in IE linguistics. However, this is not an introduction to historical linguistics, phonology, morphology, or syntax in general. It is assumed that the potential readers will already be familiar with basic linguistics, although the aim was to make the volume as reader-friendly as possible. The approach in the volume is mostly historical, as traditional in IE linguistics, and the volume is linguistics only – i.e., there are no chapters on IE culture, archeology, mythology, poetics, etc. In the last twenty years, a lot of historical-comparative grammars of IE in English have appeared (cf. the literature in the first three chapters), but usually by one author surveying all the languages – here, specialists in the field deal with one IE branch (or topic) only from a wider IE perspective. The chapters of the volume are meant to be complementary, but each chapter is supposed to be able to stand for itself (obviously, there is some repetition of certain facts).

The editor has tried to make this volume more uniform (e.g., the same PIE “orthography” is used throughout the book) and internally coherent than the first one, but it was impossible to “tame” completely so many authors, each with their own writing style and

varying interests. Obviously, each chapter has the author's personal touch (which is not bad – otherwise, all the introductions to IE linguistics would be the same). Naturally, not all literature and different views that would deserve it could be included in this volume. Also, the reader will sometimes note the conflicting views in the volume – not only was this impossible to avoid, but there was no reason to avoid it. This is something that the reader has to take into account when he or she encounters differing PIE reconstructions or interpretations of certain processes or phenomena. Rather than being a hindrance, this can allow a beginner to understand that historical linguistics is not a dogma, and that by applying its methods we can sometimes get differing results. It is exactly after the competition of different ideas that *communis opinio* is refined. In the end, the editor is glad to note that, although the basic idea of the volume is to provide an introduction and overview of IE linguistics and in spite of the fact that people have been doing IE historical linguistics for quite some time now, the volume does indeed contain quite a few new insights and interpretations, which will make it an interesting read not only for beginners and non-experts but also for professional IE linguists.

Finally, I would like to thank the publisher (especially the editor Samantha Vale Noya) for the opportunity to do this (though I have often, in times of despair, regretted the fact that I accepted the very difficult and often tedious editorial work), and all the authors who have worked hard in order for us to be able to produce such a wonderful volume.

REFERENCES

- Ramat, Anna Giacalone & Ramat, Paolo (eds.) 1993 [1997], *Le lingue indoeuropee*, II
Mulino, Bologna
- Ramat, Anna Giacalone & Ramat, Paolo (eds.) 1998, *The Indo-European Languages*,
Routledge, London–New York

ABBREVIATIONS

LANGUAGE/DIALECT ABBREVIATIONS

AC	Arcado-Cypriot
Aeol.	Aeolic
AI	Attic-Ionic
Alb.	Albanian
Anat.	Anatolian
Arc.	Arcadian
Arm.	Armenian
Att.	Attic
Av.	Avestan
AV	Atharva-Veda
Balt.	Baltic
BCMS	Bosnian/Croatian/Montenegrin/Serbian (= Serbo-Croatian)
Boe.	Boeotian
Bret.	Breton
BRuss.	Belorussian/Belarusian
BSl.	Balto-Slavic
Bulg.	Bulgarian
Čak.	Čakavian
CBalt.	Common Baltic
Celt.	Celtic
Clb.	Celtiberian
Class. Lat.	Classical Latin
CLuw.	Cuneiform Luwian
Corn.	Cornish
Cret.	Cretan
Croat.	Croatian
CS	Church Slavic
CSl.	Common Slavic
Cypr.	Cypriot
Cz.	Czech
Dac.	Dacian
Dan.	Danish
Dor.	Doric
E	East . . .
EBalt.	East Baltic
EGr.	East Greek
ELith.	East Lithuanian
Eng.	English

Est.	Estonian
Fal.	Faliscan
Finn.	Finnish
Fr.	French
Gaul.	Gaulish
GAv.	Gatha-Avestan (= Old Avestan)
Germ.	German
Gl.	Glossing
Gmc	Germanic
Goth.	Gothic
Gr.	Greek
Hitt.	Hittite
HLuw.	Hieroglyphic Luwian
Hom.	Homeric (Homer)
IE	Indo-European
IIr.	Indo-Iranian
Illyr.	Illyrian
Ion.	Ionic
Iran.	Iranian
It.	Italic
Ital.	Italian
Khot.	Khotanese
Lat.	Latin
Latv.	Latvian
LAv.	Later Avestan (= Young Avestan)
LCSl.	Late Church Slavic
Lep.	Lepontic
Lesb.	Lesbian
Lith.	Lithuanian
LSorb.	Lower Sorbian
Luw.	Luwian
Lyc.	Lycian
Lyd.	Lydian
M	Middle . . .
Mac.	Macedonian (Slavic)
ManMPers	Manichean Middle Persian
MBret.	Middle Breton
MDutch	Middle Dutch
Messap.	Messapian
MHG	Middle High German
MIran.	Middle Iranian
MIr.	Middle Irish
MParth	Middle Parthian
MPers.	Middle Persian
MW	Middle Welsh
Myc.	Mycenaean
N	North . . .
NE	North-East
Norw.	Norwegian

NPhryg.	New Phrygian
NW	North-West
NWGr.	North-West Greek
O	Old . . .
OAv.	Old Avestan (= Gatha-Avestan)
OBret.	Old Breton
OCS	Old Church Slavic
OCz.	Old Czech
OEng.	Old English
OHG	Old High German
OHitt.	Old Hittite
OIA	Old Indo-Aryan/Indic/Indian
OIc.	Old Icelandic
OInd.	Old Indic/Indo-Aryan/Indian (Vedic & Sanskrit)
OIr.	Old Irish
OLat.	Old Latin
OLatv.	Old Latvian
OLith.	Old Lithuanian
ONor.	Old Norse
OPers.	Old Persian
OPhyrg.	Old Phrygian
OPruss.	Old Prussian
ORuss.	Old Russian
OS	Old Saxon
Osc.	Oscan
OW	Old Welsh
OWCB	Old Welsh, Old Cornish and Old Breton
P	Proto . . .
Pal.	Palaic
Pamph.	Pamphylian
PAnat.	Proto-Anatolian
PBalt.	Proto-Baltic
PBSl.	Proto-Balto-Slavic
PCelt.	Proto-Celtic
Pers.	Persian
PGmc	Proto-Germanic
Phryg.	Phrygian
PIE	Proto-Indo-European
PIIr.	Proto-Indo-Iranian
Plb.	Polabian
Pol.	Polish
Pr.	Primitive
PSl.	Proto-Slavic
PToch.	Proto-Tocharian
Rom.	Romance
Russ.	Russian
RussCS	Russian Church Slavic
RV	Ṛg-Veda
S	South . . .

SCr.	Serbo-Croatian (= Bosnian/Croatian/Montenegrin/Serbian)
SE	South-East
Skr.	Sanskrit
Slav.	Slavic
Slk.	Slovak
Sln.	Slovene
Sogd.	Sogdian
Sp.	Spanish
Srb.	Serbian
Sw.	Swedish
Štok.	Štokavian
Thess.	Thessalian
Toch.	Tocharian (A & B)
Ukr.	Ukrainian
ULSorb.	Upper/Lower Sorbian
Umbr.	Umbrian
USorb.	Upper Sorbian
Ved.	Vedic
Ven.	Venetic
W	Welsh
W	West . . .
WGr.	West Greek
WSlav.	Western Slavic
YAv.	Young Avestan (= Late Avestan)

OTHER ABBREVIATIONS

1 sg./pl.	first person singular/plural
2 sg./pl.	second person singular/plural
3 sg./pl.	third person singular/plural
abl.	ablative
acc.	accusative
act.	active
adj.	adjective
adm.	admirative
adv.	adverb
all.	allative
aor.	aorist
a.p.	accentual paradigm
arch.	archaic
caus.	causative
coll.	collective
comm.	common (gender)
comp.	comparative
cond.	conditional
dat.	dative
def.	definite
det.	determinate

dial.	dialect(al)
dim.	diminutive
du.	dual
encl.	enclitic
f.	feminine
fut.	future
gen.	genitive
imp.	imperative
imperf.	imperfective
impf.	imperfect
ind.	indicative
indef.	indefinite
indet.	indeterminate
inf.	infinitive
inj.	injunctive
instr.	instrumental
interrog.	interrogative
intr.	intransitive
IPA	International Phonetic Alphabet
loc.	locative
m.	masculine
med.	medial (= middle)
med.-pass.	medio-passive
mid.	middle (= medial)
n.	neuter
nom.	nominative
opt.	optative
pass.	passive
perf.	perfect
perfect.	perfective
pers.	person(al)
pl.	plural
pluperf.	pluperfect
pres.	present
pret.	preterite
pron.	pronoun
prv.	preverb
ptcl.	particle
ptcp.	participle
sg.	singular
sigm.	sigmatic
st.	stem
subj.	subjunctive
suff.	suffix
sup.	supine
them.	thematic
trans.	transitive
voc.	vocative



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

INDO-EUROPEAN LANGUAGES – INTRODUCTION

Mate Kapović

Indo-European languages were originally, a couple of millennia ago, spoken from Europe to India. Now, in postcolonial times, they are spoken all over the world. English is thus, besides being a global *lingua franca*, spoken in North America, Australia, parts of Africa, and elsewhere; Spanish in Central and South America (Portuguese in Brazil); French in parts of Africa, etc. According to *Ethnologue*,¹ 437 IE languages (the number is in reality approximate) are spoken by almost 3 billion people in the world. With regard to the number of speakers, this makes the IE language family the biggest in the world. Concerning the number of individual languages, IE is not the largest language family – for instance, the African Niger-Congo family with more than 1500 languages is much larger. In the world, there are around 7000 languages (this is not an exact number, since it is extremely difficult to define linguistically what a language actually is), so in this sense IE languages are less important, although IE is still one of the biggest language families. Historical reasons – primarily the concentration of economic power in Europe and later in North America, together with the culture, science, etc., that goes with it – make IE the best-known language family in the world.

In the territories from Europe to India, where IE languages were spoken already a few thousand years ago, not all languages are IE. Most of the languages spoken in Europe today are IE, the exceptions (disregarding Balkan Turkish, the languages of the Caucasus, and recent migrant languages) being Basque, Maltese (an offspring of Arabic), and the Uralic languages (Hungarian, Estonian, Finnish, and a few smaller ones). Basque is the only remnant of pre-IE European languages. Other such cases (like Etruscan) have long disappeared. In the Middle East, we find both IE (like Persian/Farsi, Pashto, Kurdish, and Tajik) and non-IE languages (like Turkish, Azeri, Arabic, etc.). In the Indian subcontinent, many languages are IE (like Hindi/Urdu, Bengali, Punjabi, and Nepali), while many belong to the non-IE Dravidian family (like Tamil and Malayalam) and other non-IE language families (like smaller languages from the Tibeto-Burman and Austroasiatic families).

What is a language family? What does it mean when we say that English and Albanian belong to the IE language family? To put it succinctly, a language family consists of all the languages that have evolved from a single original proto-language. It is thus a group of genetically related languages (up to a certain point – it just may well be that all human languages are related, though this cannot be proven). In the IE case, it means that we can prove that all the IE languages that we know have evolved through many stages from a single parent language that we today call Proto-Indo-European. This proto-language was



MAP 1.1 IE LANGUAGES TODAY

Source: Adapted from C. Watkins *The Indo-European Languages* (Routledge) 1st ed (1998)

spoken some 6000 years ago and, like most proto-languages, is unattested (which means that we have no written records of PIE). This is not always the case; for instance, Latin (as the proto-language of all Romance languages) and Old Chinese (as the proto-language of all modern Chinese “dialects”/languages) are attested.

How can one language split into more than 400 languages? The answer is simple – through language change. All languages change all the time (even today with all our schools, means of communication, and mass media). Linguistic changes that occur during the life of an individual speaker are usually slow and gradual (although still perceivable, whether they have to do with changes in pronunciation, grammatical forms, or vocabulary), but in time languages change enough to become mutually incomprehensible and sometimes very different. This is made easy by migration – when two social groups separate, it is easy to imagine that a previously uniform language can evolve in different directions (cf. the differences between British, American, and Australian English, although they are still in contact, mutually comprehensible, and thus considered one language).

That languages like English, Czech, and Welsh have all developed from a single proto-language was not always apparent. Already in ancient times, people noticed that there are many similar words in, for instance, Latin and Greek – cf. e.g. Lat. *māter* ‘mother’, *pater* ‘father’ and Gr. μήτηρ /mētēr/, πατήρ /patēr/ (or, for that matter, English *mother* and *father*). However, they did not know how to properly explain such similarities. Thus, there have been many unscientific hypotheses, such as that all languages are derived from Hebrew (considered a holy language) and so on. In the 15th and 16th centuries, there were a number of scholars, like Rodolphus Agricola/Roelof Huisman (1443/4–1485), Sigismund Gelenius/Zikmund Hrubý z Jelení (1497–1554), and Johann Elichmann (1601/2–1639), who noted correspondences between various IE languages. However, it is Marcus van Boxhorn (1612–1653), a professor at Leiden University, who

can be regarded as the first historical linguist and the father of IE linguistics. In 1647 he advanced a theory that Greek, Latin, Persian, Old Saxon, Dutch, German, Gothic, Russian, Danish, Swedish, Lithuanian, Czech, Croatian, and Welsh (but not Hebrew) originally stemmed from “Scythian” (= PIE). His friend Claudius Salmasius/Claude de Saumaise (1588–1653) added Sanskrit to the list. Amazingly, van Boxhorn also noted the importance of recognizing false cognates, loanwords, systematic correspondences, plausible semantic agreement, morphological comparison, and synchronically irregular forms (all basic elements of historical linguistics now called the “comparative method”).² The European discovery of Sanskrit played an important role in the beginnings of IE historical and comparative linguistics. Already in 1585, Filippo Sassetti (1540–1588), a merchant and scholar, noted in a private letter the similarity of some words in Sanskrit and Italian (e.g., Skr. *nava* and Ital. *nove* ‘nine’), and in the 18th century a couple of scholars wrote on the obvious similarity of Sanskrit to European languages. One of them was William Jones (1746–1794), who in a lecture in 1786 compared Sanskrit to Greek, Latin, Gothic, Celtic, and Old Persian, claiming that they “have sprung from some common source.” Jones is usually wrongly considered the founding father of IE historical linguistics, although he was neither the first to make such a hypothesis nor the one with the clearest presentation of the problem. Shortly after, the first modern IE historical linguists – Jacob Grimm (1785–1863; famed as one of the Grimm brothers), Rasmus Rask (1787–1832), and Franz Bopp (1791–1867) – started doing serious historical linguistics, and others followed. The very term *Indo-European* was coined in 1813 by Thomas Young (1773–1829). The discovery of Anatolian and Tocharian texts in the beginning of the 20th century gave new impetus for IE linguistics.

The IE language family comprises ten principal branches (those being the ones sufficiently attested in ancient times and/or today) and a number of fragmentarily attested separate languages spoken in ancient times (like Lusitanian or Phrygian) that have since disappeared. Three main branches (Anatolian, Indo-Iranian, and Greek) are attested already in the second millennium BCE, two in the first millennium BCE (Italic, Celtic), and the other five branches later. Three of the principal branches consist of one language only (Greek, Armenian, and Albanian), although those comprise different dialects as well. Two major branches (Anatolian and Tocharian) are now extinct, and the same goes for all fragmentarily attested languages and certain languages from other primary branches (like Gaulish). The ten principal IE branches are as follows.

1. **Anatolian** – attested from the 19th/16th to the 1st century BCE

Spoken, until its disappearance, in Asia Minor. Old Anatolian languages (spoken in the 2nd millennium BCE) are Hittite (the best attested, written in Akkadian cuneiform, the language of the great Hittite Empire), Cuneiform Luwian (Luvian), Hieroglyphic Luwian, and Palaic. Hittite is attested from around the 16th century, but a few names and loanwords are found already in the 19th century BCE in Old Assyrian texts. New Anatolian languages (spoken in the 1st millennium BCE) are Lydian, Lycian (Lycian A), Milyan (Lycian B), Carian, Sidetic, and Pisidian.

2. **Indo-Iranian** – attested from the 15th–14th century BCE

Spoken mostly in the northern part of the Indian subcontinent and the Middle East (Eastern Turkey, Caucasus, Iran, Afghanistan). As a result of old migrations, the Romani language dialects are spoken by the Roma people in Europe. Indo-Iranian consists of two great groups – Indo-Aryan (Indic) and Iranian. The third, smaller group is Nuristani (on the border between Afghanistan and Pakistan), and the fourth, disputably, is Dardic (traditionally considered part of Indo-Aryan). The two major

groups are attested from ancient times and represented by many languages, while the two small groups are attested only from modern times and consist of a few small languages spoken in remote parts of northern India, Afghanistan, and Pakistan. Indo-Aryan languages are chronologically divided into Old Indic/Indo-Aryan (Vedic – the language of the Vedas; Sanskrit – the classical language of India), Middle Indic/Prakrits (Pali, Shauraseni, Maharashtri, Magadhi, etc.), and New Indic (many languages – Hindi/Urdu, Bengali, Punjabi, Marathi, etc.). Certain words and names of Indo-Aryan origin were attested in the 15th–14th century BCE in Hurrian texts. The Rigveda (the oldest part of the Vedas) is usually dated to the second half of the 2nd millennium BCE but was transmitted orally for a long time. Iranian languages are chronologically divided into Old Iranian (Avestan – the language of Zarathustra, from the first half of the 1st millennium BCE, or perhaps even earlier but at first transmitted orally; Old Persian – the language of the Achaemenid Empire), Middle Iranian (Pahlavi – the language of the Sassanid Empire; Sogdian; etc.), and New Iranian (New Persian/Farsi – the official language of Iran; the Kurdish languages; Pashto – the official language of Afghanistan; etc.).

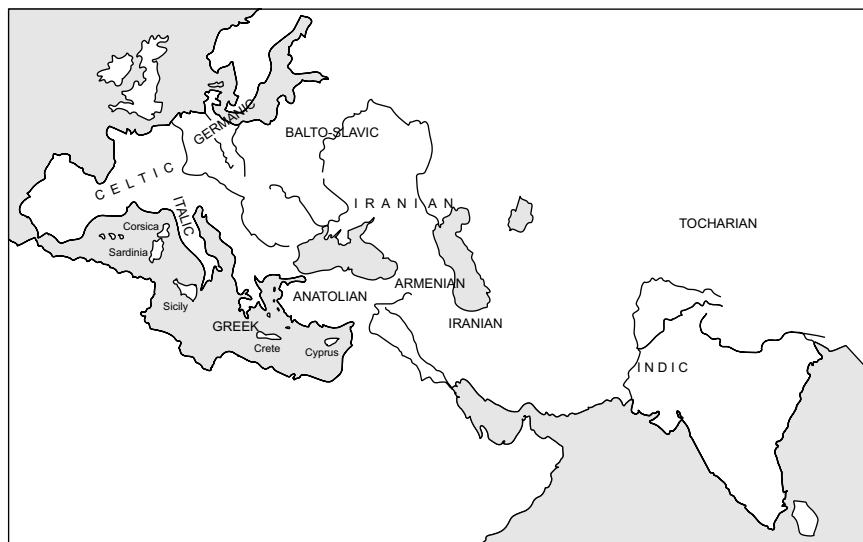
3. **Greek** – attested from the 15th–14th century BCE
Spoken from ancient times until the present in mainland Greece and neighboring islands. The first attestations were written in the Linear B script in Mycenaean Greek, the earliest recorded Greek dialect. From the 8th century, Greek has been written in the Greek alphabet. In the classical period (around the 5th century BCE), the following Greek dialects existed – North-West dialects, Doric, Aeolic (controversial), Arcado-Cypriot, and Ionic-Attic (the details of the subgrouping are disputed). Later stages of Greek are Hellenistic/Koiné, Byzantine/Medieval, and Modern Greek.
4. **Italic** – attested from the 7th century BCE
Spoken in ancient Italy. Italic consists of two groups – Latino-Faliscan and Sabellian (Osco-Umbrian). The Latino-Faliscan group consists of Latin (very well attested) and Faliscan. The Sabellian group consists of Oscan, Umbrian, and a number of smaller and poorly attested Italic languages. All Italic languages except Latin disappeared already in Roman times owing to the expansion of Latin as the language of the Roman state. Latin later developed into numerous modern Romance languages (Portuguese, Spanish, Catalan, French, Italian, Romanian, etc.).
5. **Celtic** – attested from the 6th century BCE
Now spoken in peripheral areas of the British Isles (Ireland, Wales, Scotland) and France (Brittany), previously spoken in wide areas of Europe (from the Iberian Peninsula and France all the way to Asia Minor). The Celtic branch can be divided (though controversially) into three groups: Celtiberian, Continental Celtic, and Insular Celtic. Celtiberian is a one-language group, attested in the 2nd to 1st century BCE in present-day Spain. Continental Celtic languages – Gaulish (in France) and Lepontic (in northern Italy) – are now extinct. Insular Celtic consists of Brythonic (Welsh, Breton, and extinct Cornish) and Goidelic (Irish/Gaelic, Scottish Gaelic, and extinct Manx). Of all currently existing Celtic languages, only Welsh has a more or less secure future.
6. **Germanic** – attested from the 3rd (?) century BCE
Nowadays spoken in North-West Europe. The first Germanic words are attested perhaps already in the 3rd (?) century BCE on the so-called Negau helmet (found in

Ženjak in present-day Slovenia). The earliest runic inscriptions date from the first centuries CE, and the Gothic translation of the Bible from the 4th century. Germanic languages comprise three groups: East, North, and West. The extinct East Germanic group consists of Gothic, Vandal, and Burgund. The Northern languages are Old Norse and, in modern times, the Scandinavian languages. The Western languages are Old High German, Old Saxonian, Old Low Franconian (Old Dutch), Old Frisian, and Old English (in modern times, respectively, German, Low German, Dutch, Frisian, and English).

7. **Armenian** – attested from the 5th century CE
Spoken in present-day Armenia and elsewhere in the Caucasus and Middle East as a minority language, previously also in wide areas of present-day eastern Turkey. Modern Armenian consist of two dialects, Eastern and Western, both standardized.
8. **Tocharian** – attested from the 4th–5th to the 8th–10th century CE
Spoken in merchant cities on the ancient Silk Road, in the present-day Xinjiang province of western China (now inhabited by Turkic Uyghurs). There were two closely related Tocharian languages – Tocharian A and Tocharian B.
9. **Balto-Slavic** – attested from the 9th (Slavic) and 16th (Baltic) century CE
Spoken in the Baltic, the Balkans, and Central and Eastern Europe. The Balto-Slavic branch consists of West Baltic, East Baltic, and Slavic. The extinct Old Prussian is West Baltic, and Lithuanian and Latvian are East Baltic (there are a few other poorly attested and extinct Baltic languages). Slavic consists of East (Russian, Belorussian, Ukrainian), West (Polish, Czech, Slovak, etc.), and South Slavic (Slovene; Bosnian/Croatian/Montenegrin/Serbian – traditionally known as Serbo-Croatian; Macedonian; Bulgarian). The first attested Slavic language, Old Church Slavic, was based on the Thessaloniki Old Macedonian dialect.
10. **Albanian** – attested from the 15th century CE
Spoken in present-day Albania, Kosovo, West Macedonia, South Montenegro, North-West Greece, and South Italy. There are two dialects – Gheg in the north and Tosk in the south, the latter being the basis of Standard Albanian.

Fragmentarily attested IE languages can be divided into two basic groups: those that have at least one attested text/inscription and those that are attested only through onomastics or individual words in texts written in other languages. In the first group, there are Lusitanian (Portugal), Venetic (North-East Italy), Messapic (the “heel” of Italy), Thracian (Bulgaria), and Phrygian (Minor Asia). In the second group, we find languages like Ligurian (North-West Italy), the Illyrian languages (North-West Balkans), Dacian (Romania), and Macedonian (no relation to present-day Slavic Macedonian). Apart from fragmentarily attested IE languages, there are also possible non-attested IE languages, whose traces can perhaps be seen in loanwords in other languages (like the supposed North-West Block language in Benelux), but these are rather speculative.

Many elements can be reconstructed for the last stage of PIE. However, what we obtain through reconstruction will never be a full picture of the language, since elements (sounds, grammatical endings, words, meanings, etc.) can disappear with no trace or leave so few traces that it is not possible to reconstruct them plausibly. Thus, the around 1500 PIE lexical items that have been reconstructed are certainly just a small part of the actual PIE vocabulary, which must have had something like 30,000–50,000 words, just



MAP 1.2 IE LANGUAGES IN THE 1ST MILLENNIUM BCE

Source: Adapted from C. Watkins *The Indo-European Languages* (Routledge) 1st ed (1998)

as much as any language today.³ Likewise, we can, for instance, reconstruct the ending variants *-mes, *-mos, *-mesi, *-men, etc. in the PIE 1st person plural present indicative (p. 93), but it is impossible to know how to interpret them exactly. Were these just random variants (not likely), were they sociolinguistic variables, did their use depend on dialect, gender, social status, etc.? Sometimes it is also very difficult to separate different layers – should an element be reconstructed for PIE in general, dialectal PIE, post-Anatolian PIE (i.e., PIE after the early split of Anatolian), early PIE (i.e., PIE including Anatolian), post-PIE time, etc.?

Like most languages that cover any substantial ground, PIE probably also had dialects. The dialectal picture of PIE is not completely clear because it is sometimes difficult to distinguish PIE dialectal features from post-PIE shared innovations and accidental independent innovations. It is now widely accepted that Anatolian was the first branch to separate from the rest of PIE (and that Tocharian was the next). However, which Anatolian characteristics (like the lack of feminine gender) are archaisms, and which are innovations, is still a matter of dispute (p. 62, 66, 159, 178). Dialects can share one trait with one dialect, and another with another dialect. For instance, Greek, Indo-Iranian, and Armenian all have the augment *h₁e- in past tense verbal forms, which is not present in other IE languages. At the same time, Armenian and Indo-Iranian are satem languages, while Greek is a centum language (p. 27). In Greek, Armenian, and Phrygian, initial pre-consonantal laryngeals (*H-) yield vowels (p. 48), while this is not the case in Indo-Iranian. Indo-Iranian, on the other hand, shares the so-called RUKI rule (p. 206) with Balto-Slavic (p. 29, 495, 526–527), a satem branch, while Balto-Slavic shares the ending *-mos in the dative plural with Germanic (as opposed to *-b^hos in Italic, p. 63), a centum branch, etc.

The question of the PIE homeland (the “original” territory where speakers of PIE lived) is still a subject of debate. The most commonly accepted hypothesis in the last few decades has been the “Kurgan hypothesis” (Kurgans being a type of burial mound), which

puts Proto-Indo-Europeans in the steppes of present-day Ukraine and South Russia, north of the Black Sea. The position of the PIE homeland is often deduced from certain linguistic indications, and by trying to connect linguistic evidence with archeology, paleobotany, and other disciplines. The way linguistic evidence is used in trying to deduce the place of the PIE homeland follows a certain logic. For instance, since one can reconstruct a PIE word for ‘snow’ (**snoyg*^{whs}), we can suppose that PIE was spoken in a place where there was snow (which by itself does not amount to much). Evidence has to be carefully evaluated – for instance, Lat. *mare*, OCS *morje*, and OIr. *muir* ‘sea’ would all point to PIE **mori* ‘sea’ (and thus to a PIE homeland by the sea). However, Goth. *marei* (cf. Eng. *mere*, *marsh*) means both ‘sea’ and ‘lake’, which brings the PIE meaning ‘sea’ in doubt (it is not unimaginable that the original meaning was ‘lake, standing water’ and that the meaning ‘sea’ is secondary, though there can be no certainty in any of the possibilities).

Comparative linguistics has proven beyond reasonable doubt that all IE languages stem from the same proto-language. However, the question of a deeper genetic relationship of IE with other language families is still a matter of heated debate (though most IE-ists, curiously, ignore the problem, choosing not to go outside of inner-IE problems). The most likely IE relative is Uralic. There are some striking cognate-looking examples, cf. e.g. Proto-Uralic **weti* ‘water’ and PIE **wodʰ* (this is not a word easily loaned), or Proto-Uralic verbal endings 1 sg. **-mi*, 1 pl. **-me*, 2 pl. **-te*, which are practically identical to PIE **-m(i)*, **-me*, **-te* (p. 93). Although the Indo-Uralic hypothesis is still not strictly proven, grammatical correspondences like these make this proposition quite probable. On a wider scale, IE often connected with “Nostratic” languages, which include IE, Uralic, Kartvelian (like Georgian), Altaic (a suspicious family by itself, consisting of Turkic, Mongolic, Tungusic, and possibly other languages), Dravidian, Afro-Asiatic (consisting of Semitic, Egyptian, etc.), and possibly other families/languages (the “cast” is not always the same). There is a lot of work left to be done in long range comparison, but it is not at all certain that we will ever get satisfactory answers to these problems. Many linguists are, with good reason, highly skeptical concerning the very methodological basis of such reconstruction, and simply do not believe that there is enough linguistic material preserved (since sounds, forms and words not only change but can also disappear leaving no trace) to let us plausibly reconstruct proto-languages so distantly in the past.

Up till now, we have seen examples of cognate words and reconstructed PIE forms (marked with an asterisk to indicate them not being attested). But how do linguists actually know what words/forms to compare, and how do they get the idea that two or more languages are genetically related in the first place? Looking for similar words with similar meanings is obviously not enough, since similar words with similar meanings in two languages can easily be the result of linguistic borrowing. Compare Modern English and French pairs like *city/cité*, *people/people*, or *part/part*. These are all French loanwords in English, which can be proven by historical data (French loanwords were brought to England by French-speaking Normans after their 1066 conquest). When looking for genetically related languages, it is safer to look for similarities in grammar (morphology, i.e., nominal or verbal endings, etc.), since morphological elements are borrowed much less frequently than vocabulary. Looking for morphological similarities may be a problem in languages which are not rich with morphology (like Chinese or even English), but not in older IE languages, which have quite complex morphology. Some of the similarities, clearly seen in ancient IE languages, are visible even today. For instance, compare the plural ending *-i* in Italian *lupi* ‘wolves’ and Russian *волки* /*vólki*/. However, one does need to reconstruct vocabulary as well. When looking for old words inherited directly from PIE, one has to find words that are likely to have been in a language already a

couple of millennia ago (thus, not words like *computer* or *tennis*) and words that are not so likely to be borrowed. Although in principle any word can be borrowed (and all kinds of examples from all around the world can be found), in some types of vocabulary one often finds old inherited words, for instance, in words for family relations (cf. Eng. *father*, *mother*, *brother* and related Ital. *padre*, *madre*, *fratello*), numbers (cf. Eng. *two*, *three*, *seven* and related Ital. *due*, *tre*, *sette*), body parts (Eng. *nose*, *beard*, *ear* and Ital. *naso*, *barba*, *orecchio*), etc. Genetically related words are sometimes obviously similar (Eng. *mother* ~ Ital. *madre*), sometimes less obviously so (Eng. *wolf* ~ Ital. *lupo* – the *l* is the same, and the sounds *f* and *p* are phonetically close), and sometimes almost completely different (Eng. *hornet* ~ Ital. *calabrone*). Of course, similarity of form and meaning may be completely accidental, cf. e.g. non-related pairs such as Eng. *much* and Span. *mucho*, Eng. *iron* and Span. *hierro*, Eng. *day* and Span. *día*, etc. (this is more apparent when one considers phonologically more archaic Germanic cognates such as Scots *mickle*, German *Eisen*, or Swedish *dag* respectively, or the Latin forms *multum*, *ferrum*, *diēs* that yield the cited Spanish forms).

But similarity alone is not enough. Not only does it not guarantee that two words are related, it is also not scientific. What historical linguists look for are not necessarily similar sounds, words, or forms, but systematic sound correspondences. This means that certain phonemes (or clusters of them) in purportedly genetically related languages have to systematically correspond to each other, whether they are phonetically close or not. For instance, in initial position we find in English and Italian the correspondences *n-* ~ *n-* (*no/no*, *night/notte*, *nine/nove*), *h-* ~ *c-* (*hornet/calabrone*, *hound/cane*, *heart/cuore*), *t-* ~ *d-* (*two/due*, *ten/dieci*, *tooth/dente*), etc. Careful examination and analysis of such correspondences is one of the primary tasks of historical linguistics and a prerequisite for proving a genetic relationship between two or more languages. Describing such correspondences will be an important part of the next chapter.

Note: I would like to thank David Mandić and Petra Šoštarić for reading the first draft of the chapter. Of course, all the mistakes are just mine.

FURTHER READING

Lehmann 1967 provides a rich selection of translated primary texts (with commentary) from the beginnings and early IE historical linguistics. Pedersen 1959 is also still very useful. For the problems of the IE homeland, culture, and archeology from a mainstream IE-ist perspective see, for instance, Mallory 1989 and Mallory & Adams 1997. A lot of information on IE culture can be found in the famous but controversial and idiosyncratic Gamkrelidze & Ivanov 1995. For more on the “Kurgan hypothesis” from the pen of its originator see Gimbutas 1997. Salmons & Joseph 1998 is an informative look at Nostratic from different perspectives. There are a number of good handbooks and introductions to historical linguistics, many of them using a lot of IE material. Trask 2007 may be a good place to start, as well as Campbell 2004. Joseph & Janda 2003 and Ringe & Eska 2013 are a bit more demanding. Labov’s trilogy on the principles of linguistic change (1994, 2001, 2010), written mostly (but not exclusively) from the perspective of quantitative sociolinguistics, is invaluable and inspiring, though perhaps not a usual IE-ist read. For a list of recent introductions to IE, see the “Further reading” sections in the chapters on PIE phonology and PIE morphology in this volume.

NOTES

- 1 <http://www.ethnologue.com/statistics/family> [last approached: Feb. 2, 2016].
- 2 For van Boixhorn and other early scholars cf. Van Driem 2001: 1039–1051.
- 3 Cf. Mallory & Adams 2006: 117–119.

REFERENCES

- Campbell, Lyle 2004, *Historical Linguistics. An Introduction*, Edinburgh University Press, Edinburgh/MIT Press, Cambridge
- Gamkrelidze, Thomas V. & Ivanov, Vjačeslav V. 1995, *Indo-European and the Indo-Europeans. A Reconstruction and Historical Analysis of a Proto-Language and a Proto-Culture*, Mouton de Gruyter, Berlin–New York
- Gimbutas, Marija 1997, *The Kurgan Culture and the Indo-Europeanization of Europe. Selected Articles from 1952 to 1993*, Robbins Dexter, Miriam & Jones-Bley, Karlene (eds.), Institute for the Study of Man, Washington, DC
- Joseph, Brian D. & Janda, Richard D. (eds.) 2003, *The Handbook of Historical Linguistics*, Blackwell Publishing, Oxford
- Labov, William 1994, *Principles of Linguistic Change. Volume 1: Internal Factors*, Blackwell, Oxford–Cambridge
- Labov, William 2001, *Principles of Linguistic Change, Volume II: Social Factors*, Blackwell, Oxford–Cambridge
- Labov, William 2010, *Principles of Linguistic Change. Volume 3: Cognitive and Cultural Factors*, Wiley-Blackwell, Oxford
- Lehmann, Winfred P. 1967, *A Reader in Nineteenth Century Historical Indo-European Linguistics*, Indiana University Press, Bloomington–London
- Mallory, James P. 1989, *In Search of the Indo-Europeans: Language, Archeology and Myth*, Thames & Hudson, London
- Mallory, James P. & Adams, Douglas Q. (eds.) 1997, *Encyclopedia of Indo-European Culture*, Fitzroy Dearborn Publishers, London–Chicago
- Mallory, James P. & Adams, Douglas Q. 2006, *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*, Oxford University Press, Oxford–New York
- Pedersen, Holger 1959, *The Discovery of Language: Linguistic Science in the 19th Century*, Indiana University Press, Bloomington
- Ringe, Don & Eska, Joseph F. 2013, *Historical Linguistics. Toward a Twenty-First Century Reintegration*, Cambridge University Press, Cambridge
- Salmons, Joseph C. & Joseph, Brian D. (eds.) 1998, *Nostratic: Sifting the Evidence*, John Benjamins, Amsterdam–Philadelphia
- Trask, Robert L. 2007², *Trask's Historical Linguistics*, revised by Robert McColl Millar, Hodder Arnold, London
- Van Driem, George 2001, *Languages of the Himalayas: An Ethnolinguistic Handbook*, vol. II, Brill, Leiden