Linguistic Perspectives on the Construction of Meaning and Knowledge

Linguistic Perspectives on the Construction of Meaning and Knowledge

Edited by

Brian Nolan and Elke Diedrichsen

Cambridge
Scholars
Publishing

Linguistic Perspectives on the Construction of Meaning and Knowledge

Edited by Brian Nolan and Elke Diedrichsen

This book first published 2019

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-3899-0
ISBN (13): 978-1-5275-3899-3

TABLE OF CONTENTS

SECTION B. SEMANTIC CHALLENGES IN DERIVING MEANING

SECTION C. COMPUTATIONAL APPROACHES TO MEANING CONSTRUCTION

SECTION D. DIGITAL ONTOLOGIES, AND THEIR ROLE IN MEANING

CHAPTER ONE

PERSPECTIVES ON THE CONSTRUCTION OF MEANING AND KNOWLEDGE

BRIAN NOLAN, ELKE DIEDRICHSEN

The **theme** of this book is the exploration of the dimensions of the construction and management of meaning in language, from several important topical perspectives that are of major interest to scholars today, in the realms of pragmatics, semantics, ontological knowledge engineering, and computational linguistics. This book brings together **researchers** from a variety of functional, cognitive, computational and knowledge engineering theoretical backgrounds who have worked on the nature of meaning in language, within one language or from a cross-linguistic perspective at the syntax–semantics–pragmatic, or computational–knowledge engineering interfaces. As such, the general **objective** is, through studying the pragmatics of language in interaction, semantics and syntax within the framework of functional, cognitive constructional, and computational approaches, to bring new focus and fresh perspectives that integrate many aspects of meaning construction to arrive at a yet better understanding of the cross linguistic behaviour of these dimensions. In this book an impressive **variety of languages** is represented, including Indo-European languages such as Irish, German, Spanish, Chilean Spanish, English, French, Russian, and also Pitjantjatjara, Yankunytjatjara and Ngaanyatjarra from Australia's Western Desert region, and Irish Sign Language.

The **purpose** of the book is therefore to draw a comprehensive, representative and detailed picture of the linguistic, pragmatic, ontological and computational dimensions of meaning, across a rich set of languages, in order to arrive at a better understanding of the nature and rich complexity of meaning. The **topics** that are discussed include: pragmatic approaches to meaning in the resolution of sentence meaning vs. utterance meaning; the ways in which context and situation play a role in meaning construction, and the role of core and emergent common ground in the management of meaning in language use.

The **questions** addressed across the book's chapters include:

1. How is context and common ground managed and constructed in human language-aware software?
2. What are the motivations and applications of Human Language Technology (avatars, digital corpora)?
3. What are the linguistically-motivated digital and computational strategies for constructing meaning?
4. How do different languages present different challenges for utterance meaning (written, spoken, sign languages, and Internet-based language use)–we live in a multilingual globalised world
5. How does emergence of meaning operate across Social Media, online political dialogue and online texts?
6. What are the critical issues in dealing with meaning in cyberbullying through linguistics and IT/text mining strategies?
7. How might insights from contemporary data analytics and statistical approaches complement linguistic strategies in construction of meaning?
8. How can Internet memes be described in terms of linguistic convention and change?
9. How do we address the challenges in identifying meaning at lexical, syntactic, pragmatic levels with insights from knowledge engineering and computational approaches to meaning?

The **aim of the book** is to provide a comprehensive exploration of the dimensions of meaning and knowledge representation within a number of important perspectives including linguistic, pragmatic, knowledge engineering and computational paradigms and analyses. Our **theoretical framework** is situated within modern functional-cognitive constructional-ontological and computational paradigms and our analyses are supported by authentic data (including corpus data) from the languages concerned. We find **evidence** that meaning construction manifests considerable variability in cross-linguistic comparisons in the construction of pragmatic common ground. Context and situation play an important but complex role in meaning elaboration. The role of context and situation is elusive and has proved difficult to elucidate with respect to meaning and knowledge representation. We find evidence on the nature of the, often rapid, emergence of meaning in the digital world of Semantic Web, Internet, Social Media, and Internet memes within a global multilingual world. In this era of global ubiquitous communications, where meaning seems to *mean what you want it to mean at a moment of time*, *whether facts are true or not*,

we deliver important insights into meaning in discourse across a number of domains, including for example, the political domain. Dependable definitions of what entities and propositions mean are essential for certain industries and domains. We provide core insights into meaning situated in digital ontologies across global domains, for example, the aerospace industry where rigorous definitions are critical for safe communications. The use of computational avatars and the rise of human language technologies, including massive digital corpora and big data, has made the construction of meaning and human language understanding essential to the work of linguists, cognitive scientists and computer scientists who are increasingly working together in collaborative teams to share insights.

The book is organised into **four sections** (with multiple chapters within each section) to address these topics in a cohesive and thematically coherent manner. These sections are as follows: A) Meaning in language in interaction–pragmatic challenges; B) Semantic challenges in deriving meaning; C) Computational approaches to meaning construction; D) Digital ontologies and their role in meaning.

The papers in section A: Meaning in language in interaction–pragmatic challenges, recognises that pragmatics is the study of language in meaningful interaction. Pragmatics is concerned with the use and meaning of an utterance rather than the sentence meaning. Pragmatics therefore helps us to understand the difference between '*what is said'* and '*what is meant'*. In the retrieval of meaning from an utterance, one needs to consider the contribution of the situation in which the discourse interaction took place, the particular speech act and its felicity conditions for successful realisation, and context. All of these dimensions contribute to the discourse interlocutors' shared knowledge in common ground in order to advance meaningful discourse.

In Chapter two, the first chapter in section A, Elke Diedrichsen in her paper 'Challenges for knowledge representation: emergence in linguistic expressions and internet memes', observes that, in modern approaches to linguistics, the relationship between signifier and signified is not believed to be something static, that is once and for all stored in the mental lexicon and shared by all speakers of a language. Rather, the concept of 'emergence' has entered the discussion of the way people create and understand linguistic items and utterances, and it seems to encompass all aspects of linguistic production and comprehension. Diedrichsen argues that, as a consequence, a dynamic approach to communicative interaction should be taken, considering that people belong to many smaller or greater peer groups at a time, and that the circumstances of life and the situational conditions of an interaction can change any time. This affects the way knowledge is shared

and applied in communicative encounters. Many scholars maintain that the "common ground", which is the knowledge shared between speakers, may but need not be shared in advance of the interaction. Besides the a priori shared knowledge called "core common ground" there is also "emergent common ground", which is knowledge that comes up as part of the interaction and is dynamically integrated by the interactants. In this chapter, Elke Diedrichsen discusses aspects of the emergence of linguistic structures, including the dynamicity and other-orientedness of communication, which entails the reference to and interactive recreation of common ground. To this end, Diedrichsen gives a short introduction to Google's recent invention, the Duplex speech assistant, and evaluates the progress that has been made towards automating natural conversation with respect to its usability in communicative situations with more or less predictable background information. Diedrichsen also analyses examples of formal and functional variation in linguistic constructions, which have been described by the term 'openness'. The term entails the aspect of variability that is given in complex linguistic structures. Using a German example, Diedrichsen shows that it also encompasses the establishment of hints towards subjectivity and speaker's attitudes in a structure. In addition to linguistic structures, Diedrichsen discusses Internet memes and argues that Internet memes can be viewed as communicative units, as they establish conventions for form and meaning in a shared culture. The conventions are recognised and elaborated on by users in participatory digital media. Diedrichsen discusses two popular memes in order to demonstrate the emergence of form and meaning with them, and the background on which users operate in order to recognise, understand and procreate the formal realisation and the semantic essence of a meme, including the pragmatic function and the sentiments it carries.

Chapter three by Conor Pyle, entitled 'Tracking of referents in the Western Desert languages of Australia', provides a Role and Reference Grammar (RRG) analysis of how discourse referents are tracked in text in Pitjantjatjara, Yankunytjatjara and Ngaanyatjarra (PYN), from the Western Desert of Australia. Role versus reference has two functions in syntax signalling the role of arguments with respect to the clause and with reference to what was said in previous clauses. Cross-linguistically, new referents are generally introduced in absolutive (S or O) roles because the A role is usually the topic and is referenced by a pronoun in the narrative, whereas the O argument is often ephemeral. In PYN, characters are introduced on first mention, thereafter pronoun clitics are used, being cognitively lighter than full pronouns: a zero 3rd person default clitic and ellipsis extend this trend, a null pronoun being a zero anaphor retaining salience from a

previous clause. This leads to verb rich utterances, with verbs frequently in series. Thus, an argument is backgrounded once it has been established in discourse, which is part of 'information flow'. PYN also has switch reference particles and sub-clauses which obviate the need for overt expression of syntactic subject. This chapter also draws on the idea of 'Common Ground', which is mutual knowledge, beliefs and assumptions. As participants speak, they 'ground' what has been said in the conversation. There is a presupposition by the speaker of what is common ground. Thus, a sentence may be appropriate only in a particular situation. Core common ground (including common sense and cultural knowledge) is distinguished from emergent common ground which builds during a conversation. In small communities there is a high degree of local knowledge so there is no need to specify everything in conversation, and cognate verbs imply the existence of an undergoer that does not need to be overtly expressed. Centering theory refers to the centre of attention in a conversation and this affects the form that referring expressions take. Forward looking centres are discourse entities evoked by an utterance, while backward looking entities are similar to topics. As conversation progresses the topics under discussion develop and change. Centering theory seeks to address anaphora resolution. There is rich information in a first utterance, but memory of utterances fades rapidly which means unless referents are constantly refreshed, they may need to be explicitly stated again. PYN arguments thus do not need to be specified; though it leaves a sentence technically incomplete; and reference crucially depends on context. These may be accounted for by exophoric expressions deriving from the situation; endophoric ones referring to something already in the text or homophoric ones deriving their interpretation from cultural reference. This study characterises how this is accounted for in PYN.

In chapter Four, 'The dominant principle of meaning construction in mind and discourse', by Nikolay N. Boldyrev, it is argued that knowledge representation and meaning construction in mind and discourse is always situated and is a cooperative event. The relationship between knowledge about the world and language use is indirect and depends on how speakers of a language define it. For Boldyrev, this issue suffers from a lack of profound insight into the conceptual aspects of verbal interaction and needs thorough consideration of the core conceptual factors governing knowledge representation and construction of meanings in mind and discourse. In this chapter, Boldyrev argues that the fundamental principle that underlies cooperative communication is the Principle of Interpretation Interaction. This involves conceptual accommodation, interpretation and negotiation of meanings within contexts of collective and individual knowledge activated

in participants' minds in discourse. This idea is that there are many ways by which individuals can construct their world. The problem of knowledge representation and meaning construction is central to current theoretical and empirical research oriented towards the study of cognitive processes and their instantiation in language. Knowledge representation and meaning construction is argued to involve three functions of language, which are cognitive, communicative and interpretive. All three are important and significant. These three functions are intended to account for the three types of knowledge representation in language: lexical representation, grammatical representation and modus representation. Correspondingly, language as a system of knowledge representation manifests itself as a threefold unity of the representative, communicative, and interpretive aspects. A cognitive theory of language is intended to deliver significant insights into the structure of human consciousness, and into the interrelations between language and mind, by providing evidence on their interdependence. There is no doubt that language as a cognitive ability is an integral part of human cognition. Therefore, Boldyrev concludes that, to achieve access to cognitive structures and processes and an understanding of how humans communicate meanings, then a deep understanding of the conceptual basis of language structure and use is necessary. A related conclusion is that people have an unbounded ability to create numerous meanings as well as new linguistic forms to represent these meanings.

In chapter five, 'The forms, functions and pragmatics of Irish polar question–answer interactions', Brian Nolan examines the challenges of unpacking meaning and characterising knowledge in the speech act of requesting information in one of its manifestations, the polar yes-no question, for Irish. Irish does not have any exact words which directly correspond to English 'yes' or 'no' and so employs different strategies where a yes-no answer is required. Nolan characterises the expressive forms, functions, logical underpinnings and pragmatics of polar yes-no interrogatives as question–answer pairs, and the felicity conditions necessary for their successful realisation. In a question-answer interaction, information is assumed to be freely exchanged, under a Gricean presumption of cooperation. A polar yes-no question in Irish can be considered as advancing a hypothesis for confirmation and consequently, there are several strategies available for answering a polar yes-no question. In Irish, the answers to yes-no questions echo the verb of the question for both affirmative and negative answers, along with a negation marker for negative answers. These types of answers are referred to as verb-echo answers. Typically, In Irish, the verb form is used *without* explicit nominal arguments expressed within grammatical relations, though there are

exceptions. Additionally, in negative polarity answers, the negative particle is also used. When a synthetic verb form is used, a pronominal appears in the grammatical relation of nominative subject within the answer. In the case of analytic verb forms, the subject is always missing. A subject is used only when the speaker chooses an emphatic affirmation or denial. The verb within the answer is inflected for tense as well as subject agreement. As tense is a clausal operator, it locates the time of the event denoted by a clause in relation to the time of utterance. The presence of tense in the answer implies the presence of a clause. When it occurs, pronominal subject marking implies the presence of a subject, hence also the presence of a clause. Under certain circumstances, as the answer to a copula-question with an indefinite predicate, the copula-derived phrases *sea* (*COP+3SG* = 'be-it') and *ní hea* (= NEG.COP *3SG* 'NEG be it'), function as logically equivalent to 'yes' and 'no'. This chapter argues towards several claims regarding polar yes-no questions of Irish. One claim is that the answers to polar yes-no questions of Irish contain instances of ellipsis and, as such, represent full clausal expressions with a complete semantics where the elided elements are from the question part of the question-answer pair. The propositional content is inferred from the context, specifically from the question with which the answer is paired. Another claim is that one of the functions of interrogatives is the maintenance of common ground via the update and exchange of information between the interlocutors. It also serves to reinforce social affiliation in a group through having access to shared knowledge and understanding. The fact that languages have clausal types for requesting information, and asking (polar yes-no) questions, shows clearly how important this activity is to human communication, and the construction and maintenance of common ground, and meaning.

Chapter six, 'Semantic Structure of the Sentence: Cognitive and Pragmatic Aspects, by Irina Ivanova-Mitsevich, starts by observing that communication may be considered as the central organizing activity in human life, and that language is the main means of communication which provides a mechanism that permits us to produce the necessary units that can transfer the information in discourse. This mechanism helps human beings to create different units each having its own function. The notion of unit is that used by stratificational grammar, and the function of the unit–a sentence–is to present information about a certain state of affairs to the speaker's discourse interlocutor. A sentence performs its function by representing the structure of some state of affairs or, in other words, a situation. Ivanova-Mitsevich argues that, usually, the sentence meaning is thought of as a replica of a situation, but that there are some logical and communicative difficulties which prevent such interpretation of the

meaning of sentences. This understanding of the meaning of a sentence can lead to the conclusion that there should be as many sentence structures as there are types of situations. However, sentence structures are not so numerous. Ivanova-Mitsevich argues that exchanging information is possible only if the participants of communication have common structures. However, a new type of situation has no structure for itself, and a structure for it should be invented by a speaker. How then might the listener interpret it since the listener does not possess the appropriate structure? Thus, the meaning of a sentence should be a result of correlation of at least two semantic structures. One of these is the structure of our knowledge about the state of affairs ("denotational field") and the other is a structure of our logics ("signification"). These two structures have different origins and functions. The linguistic aim of this chapter is to find what in the sentence semantics gives speakers the possibility to present a situation in different ways, and to find out what mechanisms underlie the variable reflection of a situation in the sentence structure. Ivanova-Mitsevich investigates the process of creating the sentence meaning, and the coordination of the necessary structures. Use of the two terms, "situation" and "denotational field", show that in the denotational sphere we have to differentiate between the cognitive structure, which is immune to syntactical structures and is a part of our cognition, i.e. the denotational field, and situation, which is a model of some denotational field that is created for communication and directly enters the semantics of the sentence. Consequently, a given denotational field might be reflected by a number of differently structured sentences, each of which presents a specific view upon the field that is its own situation. While perceiving and categorising a fragment of reality, a speaker forms a model of it, i.e. a denotational field. This model includes the most abstract ideas about the objects of the reflected fragment of reality and possible relations binding them. In the process of creating a sentence for transferring knowledge of the reflected fragment of reality to the communication interlocutors, the speaker selects a minimal number of denotational elements relevant for the communicative conditions and establishes relations among them by creating a definite point of view on the denotational field, and frames a situation. In order to present a situation in a linguistic form, the speaker has to qualify the relations existing among the components of the situation, as dynamic or static, and directed or non-directed. The speaker needs to construct an appropriate proposition to enable reflection of the situation in a sentence. In employing operations of positioning, the focus of the speaker's interest and the center of empathy, the speaker makes a logical arrangement of the nominal components of the situation.

Chapter seven, 'Linking constructions into functional linguistics: On functional-semantic characteristics of lexical-modal discourse-text 'transitions' in modern English and French', by Sabina Nedbailik, considers a discourse-text to be a coherent system, functioning as a complete message, possessing its own content and organised by abstract models, characterised by some distinctive features within the particular language. The notion of **'**content**'** of a text is different from the notions: **'**sense**'** and **'**meaning**'**. For Nedbailik, coherence and cohesion can be treated not only as semantic phenomena, being manifested simultaneously as structural, semantic and communicative integrity, **'**interacting as form, content and function**'.** As such, every text presents a regular structure, informed by a definite set of categories. The communicative integrity of a text is expressed in the relations of succession between its forming components. Each sentence is supported communicatively by a preceding one which produces various theme-rhematic chains, structuring a statement informatively from a known fact to a new one. All sentences are interlinked not only by their thematic unity and the principle of communicative progression, but also by various external signals, indicating that components form together some structural complex. The linking elements can be pronouns, articles and auxiliary verb forms, particles. Discourse-texts of different styles can also be formed by means of special elements, in the linguistics theory of logical connectives. These linking words, also called **'**transitions**'**, contribute to text structuring and facilitate reading, translating, comprehension. The connectors can guide the meaningful understanding of a discourse.

In section B: Semantic challenges in deriving meaning, the papers examine issues relating to the compositionality of semantics and meaning in the lexicon. Semantics encodes meaning with lexemes in the lexicon and when these are associated in combination, sentence meaning emerges–the *what is said*. Semantic meaning, however, is not 'just' the association of a concept with a lexeme. The construction of meaning is richer than that and may involve metaphor, metonymy and various kinds of schemata that facilitate meaning extension. Metaphor and analogy, and metonymy are powerful cognitive tools for the construction of meaning within the individual. It may also be the case that culture influences how we construct meaning, for example, in identity. The rich nature of meaning and its representation as knowledge continues to present challenges to us as linguists and scientists.

Chapter eight, starts section B. In chapter eight, 'Figurative Framing of Big Data', Inna Skrynnikov argues that the term *big data* is pervasive yet its meaning is ambiguous and confusing. By tracking the evolution of the terms "data" and "big data" the chapter reveals the meanings attached to them by

different usage communities. Skrynnikov frames the analysis of news items about big data via excerpts from the business and technology press which shows the crucial role of metaphor in conceptualising processes and phenomena of today's digital world of information. The types of metaphors employed and inferences drawn reflect and influence the perception of big data. Skrynnikov shows that the semantics of the terms "data" and "big data" have changed as a result of what she calls a confluence of social, cultural, and linguistic factors, thus replacing old meanings with new ones. Big data is known to have specific characteristics and properties that imply both the challenges and advantages of dealing with digital information. The properties of big data were initially referred to as the 3 Vs: volume, variety and velocity. Now, however, this list has been expanded to 10 properties which cover the multifaceted nature of big data. By applying metaphor to describe intricacies of the digital world, one can resolve the ambiguity and confusing surrounding notions of big data. In this regard, the explanatory power of metaphors highlights certain aspects of an unfamiliar phenomenon while obscuring other ones, thus enhancing our understanding. The role and relevance of metaphors are crucial for making data and big data meaningful and in shaping the meanings of these phenomena. A question Skrynnikov addresses is whether the interpretation of (big) data, its meaning and inferences are context- and subject-dependent. Given the advantages of the embodied approach to metaphor, this chapter follows the line of embodied cognition research in general and Lakoffian conceptual metaphor theory (CMT) in particular as applied to framing analysis of (big) data. Cognitive metaphor studies lead to applications of conceptual metaphor theory to fields such as media studies, discourse analysis, communication studies, and political science. Metaphors in political discourse enable mass media professionals to rely on their interpretative power and construct salient narratives they wish to promote. Framing pressing societal issues in a certain way through employing a set of corresponding metaphors repetitively, which in their turn evoke intended inferences, ultimately forms stable neural connections in the minds of a target audience. The transition to the information era calls for a major shift in a set of preferable metaphors we employ to make sense of digital information. Skrynnikov argues that the most effective data-related metaphors should be rooted in our embodied experience as a fundamental part of the way we think and act in the world. Skrynnikov further argues that both data and big data function in discourse as contested and evolving terms, and metaphor is a powerful and relevant cognitive mechanism for making these complex phenomena meaningful and shaping the meanings of these phenomena. The metaphors we use reflect the ways in which we view and understand big data. The chapter finds that

media discourse about big data is highly figurative, signalling our need for familiar embodied concepts to characterise the digital world. This chapter substantiates the crucial role of metaphor in conceptualising processes and phenomena of the digital world. The major claim of Skrynnikov in the chapter is that the types of metaphors employed and inferences drawn reflect and influence the perception of big data providing implications of its current conceptualisations

Chapter nine,' *Over* in radiotelephony communications' by Maria del Mar Robisco Martin recognises that, for over sixty years, aviation radiotelephony has been based on a standard phraseology designed to achieve the utmost clarity and brevity and to minimise failures in air-ground communication. It consists of codified and limited dialogues between air traffic controllers and flight crew members. The *International Civil Aviation Organization (ICAO)* created the *Proficiency Requirements in the Common English Study Group* and mandated that 1) English should be the universal medium for radiotelephony communications. 2) All pilots and controllers should pass an English language exam to achieve an ICAO operating level of competence. 3) All pilots and controllers should make global use and a correct application of the phraseology in these interactions. 4) It would be necessary to carry out studies to analyse the English language in these communications and to create teaching resources. Thus, following the ICAO's Requirements, Maria del Mar Robisco Martin, in this chapter, focuses on the language employed in radiotelephony communication, in particular, on the preposition *over*. This study is significant because the polysemy can affect the interpretation of a sentence and create misunderstandings. Prepositions are amongst the most polysemous words in English and the semantic network associated with any preposition in one language rarely overlaps with the meanings of any single linguistic form in another language and *over* is perhaps the most polysemous of the English prepositions. The aim of the chapter is to show the multiplicity and fuzziness of meaning in natural language, in contrast with the simplified view suggested by the standard phraseology. Therefore, this chapter deals with the polysemy and with the polysemous preposition *over* in particular. It is based on a corpus consisting of authentic cockpit voice recordings which have been processed using the AntConc software. The purpose is to demonstrate that, in the sampled cockpit voice recordings, *over* appears with more meanings than with the primary configuration, and to systematise the senses of *over* in this context. The findings suggest that *over* is used in eleven distinct senses and that they create a semantic network. Supporting this research, an electronic database consisting of cockpit voice recordings, belonging to an aviation accident database which includes all civil aviation

accidents of scheduled and non-scheduled passenger airliners worldwide, and which resulted in at least one fatality, was used to produce an electronic corpus of items. The recordings are taken from fatal aviation accidents which occurred between 1962 and 2002. The findings suggest that the Primary Sense is the only meaning of *over* proposed by the phraseology whereas an examination of aircraft communications shows that o*ver* is used in a range of other senses (the Primary Sense, the Covering Sense, the On-the-other-side Sense, the Transfer Sense, the Completion Sense, the More Sense, the Control Sense, the Examination Sense and the Repetition Sense) which create a semantic network.

Chapter ten, Linguistic and cognitive bases of differentiation of conceptual metaphors and metonymy, by the authors Svetlana V. Kiseleva, Nella A. Trofimova, and Irina B. Rubert, deals with the essence of metaphor and metonymy. Within cognitive science, which explores the processes of perception, categorisation and understanding of the world, metaphor and metonymy are considered as the manifestation of analogue capabilities of the human mind. The relevance of metaphor and metonymy in speech resides in the fact that these are ways to connect objective and subjective reality in order to convey to the listener not only the meaning of the statement, but also one's internal state and attitude to what was said. That is, metaphor and metonymy in speech are ways of combining our thinking with language, which allows people to communicate most effectively. These two different types of operations with signs–metaphor and metonymy–show that we need to study cognitive mechanisms. In this chapter, Kiseleva, Trofimova, and Rubert have as their goal to determine and to justify the linguistic and cognitive grounds of differentiation of conceptual metaphor and metonymy. The research reported on in this chapter consists of three main parts. The first is entitled "Metonymy in cognitive theories", and offers a review of different approaches to the study of metonymy. The second part, "Metaphor in cognitive theories", presents the theories of conceptual metaphor and the cognitive classification of metaphors, while the third part considers approaches to the distinction between conceptual metaphor and metonymy. In linguistics, metonymy is understood as a mechanism to expand the semantic range of a word. Cognitive linguistics expanded the interpretation of metonymy and separated cognitive metonymy, a mechanism for the conceptualisation of reality, from linguistic metonymy, a semantic mechanism for developing the meaning of a word. A metaphor in cognitive linguistics is understood as a mechanism, a process, and a result in a single and generalised form, a form of thinking. Specifically, if it is necessary to specify the meaning of the term *metaphor*, then the following terms are required: process, meaning, model, mechanism. The cognitive

theory of metaphor highlights its conceptual properties. It is the study of metaphor as a component of our conceptual system that determines the direction of research in modern metaphor theory. In this chapter, the authors consider metaphor in two aspects as static and dynamic, that is, as a result and as a process. Additionally, a third direction, metaphor and metonymy as two independent cognitive processes, is considered in virtue of the use of the cognitive-matrix method of research. The difference between metaphor and metonymy is seen in the fact that metaphors include a systematic projection of ontological, figurative-schematic and logical structures from the target area to the source area based on the relationship of similarity between the interacting areas. Metonymy includes the relationship of adjacency, expressed by various associative links, and leads to a referential shift. When considering the constructions of metaphorical and metonymic statements based on literal ones, the authors place emphasis on the basic action frame, because metonymic and metaphorical characteristics are associated with its certain parameters, and the action frames functioning within the matrices in metonymic and metaphorical processes are shown to have different origins.

In Chapter eleven, 'From walled off Europe to walled in identity', Natalia Iuzefovich reports on ongoing research about the "wall phenomenon" viewed as a marker of identity, at both national and individual levels. It is argued by Iuzefovich that our world has constant movement of people, not necessarily voluntarily, and people are losing the sense of identity that comes from being a member of a community. It is urgent therefore that linguists' study linguistic properties within political discourse which cause identity variation. In this chapter, Iuzefovich further argues that the most crucial changes of identities have been observed since the middle of the 20th century up to present times and these changes are revealed at large in the border wall issue. The "wall phenomenon" as seen as a marker of a human identity is very controversial and reflected mostly in political discourse. Iuzefovich argues that no study on the issue "wall and identity" can be fully integrated without a socio-cognitive perspective. The conceptual structure of "wall" is variable, dynamic, and it has been changing much due to political and socio-cultural context. Iuzefovich defends the claims that "wall off" implies an active role: we construct border walls with the intention to keep out others, 'them', strangers, enemies, etc. Being walled off makes us mentally 'walled in', we do not recognise other cultures appropriately. Limited or no communication with peoples of other cultures make US view THEM not just as 'others' but more like enemies, terrorists, and strangers. This chapter addresses several important questions: 1) What verbal means of representing the conceptual structure "wall" can be singled

out from political discourse, and how are they interrelated? 2) What promotes human identity variation? 3) What changes in intercultural and intracultural relations does the "wall phenomenon" cause?

Section C: Computational approaches to meaning construction, is concerned with the use of computational strategies for the representation of meaning in software for deployment in various language aware applications and, perhaps, over the internet within social media. Computational approaches to meaning construction need to be rigorously specific in delineating the interfaces between lexicon, semantics, morphosyntax in order to instruct the software to produce and understand grammatically correct, or even ungrammatical but acceptable, utterances. The computer scientist working with language aware software employs different strategies to natural language understanding and the extraction and construction of meaning. These often bring an NLP engineering approach to software rather than a linguist's understanding. More frequently, computer scientists collaborate with linguists towards the common goal. The strategies that computer scientists and computational linguists bring may employ complex algorithmic strategies to defining the linking system across the interfaces of human language, such as between lexicon, semantics, morphosyntax. At other times, they may employ insights from data mining and data analytics to the derivation of meaning. Computer scientists and computational linguists often build digital corpora that are machine readable for use by linguists. When computer scientists and computational linguists collaborate on the diverse languages found in the world, they typically provide very significant benefits towards our understanding of meaning and its representation, in this instance in software. Today, with the rise of important human language and cognitive technologies (such as IBM's Watson) in the second decade of the 21st century, contemporary challenges facing computer scientists and computational linguists include building software 'bots' that act as conversational agents. As language aware software, these human language technologies have enormous potential in our globalised worlds as working applications ranging from detection of cyberbullying in social media to language assistants for Deaf Sign Language users.

Aurelia Power, in chapter twelve, 'The role of previous discourse in detecting public textual cyberbullying', observes that previous work in the field of cyberbullying detection has focused solely on individual instances/posts taken in isolation, rather than part of the online conversation/dialogue. Consequently, the detection process typically considers only the information contained in the post itself, such as the presence of profane or violent words which may be indicative of cyberbullying. However, online discourse contains many instances that do

not comply with grammatical standards, or they provide insufficient information. For example, the instance *You clearly are* was labelled by annotators as cyberbullying in an authentic dataset, despite the fact that its content suggests no cyberbullying, and it was only when one considered the previous post uttered by a different user–*I am not pathetic*–that it was possible to identify one of the cyberbullying elements in the form of the offensive adjective *pathetic*. To address this limitation, Power investigates the role of previous instances/posts in identifying the missing cyberbullying elements, and she proposes a framework that relies on the definition of cyberbullying which divides information into discourse-old and discourse-new. Specifically, the focus of the chapter is on how discourse-old information is used to inform meaning, and infer some or all three necessary and sufficient cyberbullying elements: the personal marker, the dysphemistic element, and the link between them. First, Power analyses discourse-dependent instances of cyberbullying present in her dataset and proposes a taxonomy of their underlying constructions as follows: (1) fully inferable constructions, where all three cyberbullying elements, the personal marker, the dysphemistic element, and the link between them, are not explicitly present, but can be inferred from previous discourse, (2) personal marker and cyberbullying link inferable constructions where the dysphemistic element is explicitly present, but the personal marker and the link must be inferred from previous discourse, (3) dysphemistic element and cyberbullying link inferable constructions where the personal marker is explicitly present, but the dysphemistic element and the cyberbullying link are entities inferable from previous discourse, and (4) dysphemistic element inferable constructions where the personal marker and the link are explicitly present, but the dysphemistic element must be inferred from prior discourse. Power then develops resolution rules to identify the personal marker, the dysphemistic element, and/or the cyberbullying link, in other words, to transform such instances into instances that contain them explicitly, and, therefore, into instances that can be subjected to the detection rules discussed elsewhere. Importantly, these resolution rules are divided into separate sets that target: (1) polarity answers, (2) contradictory statements, (3) explicit ellipsis, (4) implicit affirmative answers, and (5) statements that use indefinite pronouns as placeholders for the dysphemistic element. Finally, algorithms are described to implement these resolution rules, using several types of information: grammatical and syntactic information, such as part of speech and dependency relations among sentential constituents, as well as pragmatic information, such as the previous posts and the user names.

Chapter thirteen, 'Detection of cyberbullying using text mining' by David Colton and Markus Hoffmann, also addresses the determination of meaning in a cyberbullying context. In this instance, it is achieved through strategies of text mining, as against lexical and linguistic means as we saw with the previous chapter. Colton and Hoffmann note that the Internet technology boom has led to a proliferation of tablets, laptops and smart phones with high-speed Internet access. This access, coupled with the advent of instant messaging, chat rooms and social media websites, has led to an Internet generation who think nothing of posting selfies, mood updates, their relationship status or anything about their life on-line. The traditional bully was the kid in school, or office worker, who got pleasure from watching their victims suffer as they verbally abused them or perhaps made fun of them or maybe even threatened them with physical violence. The bully has now moved on-line and a cyberbully now has 24-hour access to a potentially unlimited number of victims. The consequences of this cyberbullying activity are frequently read about in the newspapers, following another tragic teen suicide. To prevent this new form of bullying, it is important that technology is used to detect these cyberbullying posts. This chapter shows that the Python programming language, together with the application of text mining techniques in meaning resolution in cyberbullying detection, can be successfully applied in the automatic detection of cyberbullying text. As part of the contribution of this research, a new classified cyberbullying dataset, including detailed descriptions of the criteria used in its classification, and an in-depth analysis of several classifiers is undertaken. A novel way of determining the best overall classifier using the recall values of both the positive and negative class is suggested. Colton and Hoffmann provide an evaluation of the best models by simulating their evolution as new, previously unseen, samples are classified and then included as training data for subsequent iterations.

Irene Murtagh, in chapter fourteen, 'Motivating the computational phonological parameters of an Irish Sign Language avatar', provides an account of the computational phonological parameters of an Irish Sign Language (ISL) avatar, while motivating the phonological-morphological interface in ISL. Sign Languages like ISL are visual gestural in nature, and have no written or aural form. Therefore, in order to communicate an ISL utterance computationally, Murtagh implemented a humanoid avatar in software capable of movement within three-dimensional (3D) space. Murtagh uses the functional-cognitive RRG as the theoretical framework. Using RRG provides significant theoretical and technical challenges within both the RRG theory itself, and the software. Prior to preparing a linguistically motivated computational definition of lexicon entries to

represent ISL within the RRG lexicon, Murtagh defines the ISL phonological parameters in computational terms. In providing a definition of a linguistically motivated computational model for ISL it was necessary to refer to the various articulators (hands, fingers, eyes, eyebrows etc.), as these are what are used to articulate various phonemes, morphemes and lexemes of an utterance. Importantly, Murtagh proposes a new level of lexical representation, which describes the essential computational phonological parameters of an object as defined by the lexical item. The new level of lexical meaning proposed includes an articulatory structure level, which caters specifically for the computational linguistic phenomena of Sign Languages, such as ISL, enabling the representation of lexical items within the RRG lexicon.

Kulvinder Panesar in chapter fifteen, 'Motivating a linguistically orientated model for a conversational software agent', also uses RRG as the functional-cognitive linguistic model in her research. Specifically, Panesar proposes a linguistically orientated model of a conversational software agent (CSA) framework sensitive to natural language processing (NLP) concepts in a functional linguistic approach. She discusses the relationship between natural language processing and knowledge representation (KR), and connects this with the goals of the RRG linguistic theory in a computational implementation. Panesar proposes a design of a computational model of the RRG linking algorithm that utilises a speech act construction as a grammatical object with a sub-model of belief, desire and intention (BDI). This model has been successfully implemented in software using conceptual graphs and the resource description framework (RDF). Panesar highlights some important implementation issues that arise at the interface between language and meaningful knowledge representation.

Section D: Digital ontologies and their role in meaning, explores the development and population of digital ontologies and how meaning is defined and encoded within the ontology. This is a considerable challenge given that we live in a multi-lingual world and that, when digital ontologies are employed in software, getting the definition correct and unambiguous is crucially important. Digital ontologies therefore need to be able to make correspondences between meaningful concepts across different languages. Often, concepts have a scalar or graded nature, and the digital ontology must be able to encode the correct and appropriate level of granularity within a concept definition. Some concepts have a physical reference in the world whereas others are more ephemeral and abstract. Others are action-oriented and might be realised in language as a verb in one of its manifestations (verb, verbal noun, participle.). Often, because of the sheer size and complexity of the challenge involved in the creation of digital ontologies and capturing

meaning and representing meaning with them, domain specific ontologies are elaborated. Digital ontologies have a relationship with the lexicon and, once populated, they can be used to inform, provide interfaces to, or even generate, the lexicon of a language. A significant challenge for digital ontologies is word sense disambiguation on natural language processing. Another significant challenge for knowledge engineering with respect for digital ontologies is automatically and dynamically harvesting new knowledge from online digital sources on the Internet, and parsing the information and data found in various meaningful ways.

Chapter sixteen, 'An experimental review on methods for Word Sense Disambiguation on Natural Language Processing' by Fredy Núñez Torres, is concerned with digital ontologies and their role in meaning determination. The chapter presents a motivated proposal for reviewing and testing for the most relevant Word Sense Disambiguation (WSD) methods used nowadays on Natural Language Processing (NLP). This approach considers the development of experiments applied to a Chilean Spanish corpus that was designed based on the semantic representations available on the lexico-conceptual knowledge base FunGramKB. The chapter reports on computational procedures used for automatic WSD, such as machine learning; path-based metrics and overlapping glosses; and multinomial logistic regression. In the research reported here, a semi-automatic selection of potentially polysemous lexical units (nouns) was carried out to select instances (sentence context) for certain lexical units extracted from the written mass media corpus belonging to CODIDACH: Corpus Dinámico del Español de Chile (Dynamic Corpus of Chilean Spanish) and the selected lexical units were linked with specific concepts of the #ENTITY subontology of FunGramKB. The assembly and execution of all the experiments has been carried out using 'Data Mining Encountered' (DAMIEN), a computer environment to support linguistic research. DAMIEN integrates, in the same work environment, the different tools and techniques that can be applied in the analysis of linguistic corpora. These techniques come from different disciplines, such as corpus linguistics (e.g. frequency lists; XML processing and XSL; database administration and SQL; regular expressions; etc.), statistics (e.g. descriptive and inferential statistics; graphic representation of data; etc.), natural language processing (e.g. extraction of n-grams; derivation; morphological and syntactic analysis; POS tags; etc.), and text mining (e.g. classification and clustering).

Chapter seventeen, 'Ontology enrichment: A case study on the plants domain' by Eva M. Mestre-Mestre and Pedro Ureña Gómez-Moreno, notes that today's information technology permits vast amounts of data to quickly circulate across the Internet. Now, there is so much information available to

Internet users, that it becomes difficult to access it. Access and retrieval of this information is challenging, and there exists a 'knowledge acquisition bottleneck' and a lack of digital systems able to extract meaningful patterns from huge volumes of data. Ontologies are a preferred approach to address this problem. Ontology learning is the application of Knowledge Engineering strategies and methods to create ontologies, using automatic or semi-automatic methods, whereby the ontologies are created or increased and improved with little, if any, human intervention. Ontologies can be built up using structured, semi-structured or unstructured data. Depending on the way in which information is organised, ontologies can be based on linguistics, logics, machine learning methods, statistics-based. Linguistic approaches use different strategies, such as POS (part of speech) based patterns, semantic lexicons, or seed words. This chapter builds on a method for ontology enrichment which aimed at enlarging an ontology based on the collocational information of the features which characterise a term. The underlying reasoning is that two similarly described subordinate concepts likely correspond to the same superordinate. This motivated a strategy to locate the corresponding superordinate node for each of the terms proposed for inclusion (subordinates). In this chapter, Eva M. Mestre-Mestre and Pedro Ureña Gómez-Moreno, present the results of this experimentation in the construction of meaning via an ontology, carried out using a specialised domain related to the plant kingdom. The particularity of this domain-specific corpus is that, due to the classification of plants, the basic working units were bigrams.

The book provides a coherent and integrated set of analyses and addresses issues concerning the construction and management of meaning, and knowledge, over a diverse collection of languages from across the world in the perspectives of functional, cognitive, constructional, knowledge engineering, ontological, and computational approaches, in a range of cross-linguistic treatments. As a result, this volume represents a timely and contemporary instance of cross-linguistic comparison of these important discourse and syntax-related phenomena. Further, this volume contributes towards providing a comprehensive overview of the construction of meaning in language, which is central to our understanding of how human languages function. This includes Internet communication which represents a new level of linguistic discussion for the digital online world. This scholarly work leverages new and advanced thinking from knowledge engineering, data analytics and computer science. The book considers the contribution of context and situation to utterance meaning within a speech act in discourse, semantic meaning including lexical and compositional meaning. It also discusses the contribution of metaphor and metonymy, and

computational linguistic approaches, to the construction of meaning. Issues of knowledge representation are addressed in language aware software applications (i.e., Avatars, Social Media, and Internet Communications) at the interfaces between lexicon-semantics-syntax, and concise definitions of concept meaning in multilingual digital ontologies that motivate a machine-readable lexicon for use within human language technologies. The chapters bring fresh and relevant insights to the rich and complex dimensions of meaning and knowledge representation. The research reported on here shows how linguistic and computational strategies concerned with meaning and knowledge representation contribute to our understanding of grammar and human language, and how we use language.

The work will be of interest to the community of researchers and scholars within functional linguistics, knowledge engineers, computer scientists, and postgraduate students internationally who work with pragmatics, linguistics, semantics, knowledge engineering, data analytics, computer science, natural language understanding, human language technologies, and digital corpora, at the interfaces between pragmatics, syntax, semantics and the lexicon.

SECTION A.

MEANING IN LANGUAGE IN INTERACTION– PRAGMATIC CHALLENGES

CHAPTER TWO

CHALLENGES FOR KNOWLEDGE REPRESENTATION: EMERGENCE IN LINGUISTIC EXPRESSIONS AND INTERNET MEMES

ELKE DIEDRICHSEN

1. Introduction

In modern approaches to linguistics, the relationship between signifier and signified is not believed to be something static, that is once and for all stored in the mental lexicon and shared by all speakers of a language. Rather, the concept of 'emergence' has entered the discussion of the way people create and understand linguistic items and utterances, and it seems to encompass all aspects of linguistic production and comprehension.

The term *emergence*, established in conversation analysis and pragmatics (Hopper 1987, 1998, 2011), expresses the view that the dynamic and interactive nature of language applies to all aspects of communicative interaction. This involves the grammar, the convention of meaning of the signs used, the application and reliability of shared knowledge, the shared culture and its influence on the exchange. Rules and conventions emerge in interaction, as they are interactively created and negotiated.

Knowledge bases and many theories of linguistics seem to assume that humans acquire their language, including their grammar, their ontology and all aspects of meaning conventions once and for all, and then live and communicate with this set of linguistic knowledge items for the entire time they spend in a culture. This rather static view, however, does not consider the dynamic nature of communicative interaction, where circumstances change, people get in and out of peer groups and belong to many smaller or greater cultures at a time. Furthermore, the concept of a 'culture' and the knowledge shared in it is dynamic as well. Kecskes and colleagues (2008, 2010, 2012, 2014, Kecskes and Zhang 2009) maintain that the "common