The Evolution of Grounded Spatial Language

Michael Spranger

Computational Models of Language Evolution, No 5

language science press

Kommentarzusammenfassung für The Evolution of Grounded Spatial Language

Seite: -5

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:27:06
This layout is not the same as in other books of the series, e.g. Steels

The Evolution of Grounded Spatial Language

Spatial language and cognition have fascinated linguists, psychologists and robotics researchers for a long time. Spatial language is central to all human languages and has provided a rich background and source of extensions into other parts of language. For instance, it is often assumed that spatial language provides a source of structure for temporal language but also more abstracts domains. Importantly, spatial language is incredibly diverse across cultures. Different solutions for the spatial reference problem at the heart of spatial language range from using purely geocentric features to body parts to projective terms familiar from English.

This book presents groundbreaking robotic experiments on how and why spatial language evolves. It provides detailed explanations of the origins of spatial conceptualization strategies, spatial categories, landmark systems and spatial grammar by tracing the interplay of environmental conditions, communicative and cognitive pressures. The experiments discussed in this book go far beyond previous approaches in grounded language evolution. For the first time, agents can evolve not only particular lexical systems but also evolve complex conceptualization strategies underlying the emergence of category systems and compositional semantics. Moreover, many issues in cognitive science, ranging from perception and conceptualization to language processing, had to be dealt with to instantiate these experiments, so that this book contributes not only to the study of language evolution but to the investigation of the cognitive bases of spatial language as well.

ISBN978-3-944675-46-6

Seite: -4

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:29:12
extra space at beginning of paragraph

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:31:28
ISBN bar code?

Language Science Press
Berlin

Language Science Press
Habelschwerdter Allee 45
14195 Berlin, Germany

langsci-press.org

This title can be downloaded at:
http://langsci-press.org/

ISBN: 978-3-944675-46-6

Cover and concept of design: Ulrike Harbort

Storage and cataloguing done by FU Berlin

Freie Universität Berlin

Seite: ii

Verfasser: fkopecky Thema: Notiz Datum: 28.05.2015 13:32:32
in other books, bibliographical data is printed at this point

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:32:05
Book URL missing

mation experiments that these strategies are sufficient to allow a population to self-organize a spatial language system from scratch. The book continues by studying the origins and competition of language strategies. Different conceptual strategies are considered and studied systematically, particularly in relation to the properties of the environment, for example, whether a global landmark is available. Different linguistic strategies are studied as well, for instance, the problem of choosing a particular reference object on the scene can be solved by the invention of markers, which allows many different reference objects, or by converging to a standard single reference object, such as a global landmark.

The book demonstrates that the theory of language evolution by linguistic selection leads to operational experiments in which artificial agents self-organize semantically rich and syntactically complex language. Moreover, many issues in cognitive science, ranging from perception and conceptualization to language processing, had to be dealt with to instantiate this theory, so that this book contributes not only to the study of language evolution but to the investigation of the cognitive bases of spatial language as well.

This book would not have been possible without the hard work of the people at Sony Computer Science Laboratory Paris and the A.I. Lab at ~~the~~ Vrije Universiteit Brussels. Many of them have left traces in software and ideas that provide the background against which a book like this one becomes possible. Most notably I would like to thank the current and past members of the AI lab in Brussels and Sony CSL Paris who I have met and who have made contributions to the various software systems that underly the experiments described in this book: Katrien Beuls, Joris Bleys, Joachim De Beule, Wouter van den Broeck, Remi van Trijp and Pieter Wellens.

Martin Loetzsch and Simon Pauw had big impact on many issues discussed in this book. I am indebted to all of them for long discussions that have tremendously shaped my way of thinking and for their collaboration on different aspects of spatial language, conceptualization and embodiment.

Last but not least, I would like to thank Luc Steels who has had tremendous impact on the intellectual ideas put forth in this book, provided the necessary environment to conduct this research, and who continues to be an inspirational and visionary figure for future work.

Seite: iv

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 13:47:25

Contents

Seite: v

Verfasser: fkopecky Thema: Notiz Datum: 28.05.2015 13:50:50
General remark: Check capitalisation in Headings and this TOC

1 Introduction

Spatial language is a vast topic. This book focusses on locative phrases, which are phrases that single out objects in the physical environment with the *communicative intention* to draw attention to these objects. The following shows an example of a locative phrase from German.

(1.1)	*der*	*Block*	*rechts*	*der*	*Kiste*	*von*
	the.NOM	block.NOM	right.PREP	the.GEN	box.GEN	from.PREP
	dir	*aus*				
	your.DAT	perspective				

'The block to the right of the box from your perspective'

Phrases like this can be seen as highly complex tools that help dialog partners to establish spatial reference. The utterance conveys to the hearer a number of instructions such as (1) apply the spatial relation `right`, (2) use a particular landmark and (3) take the perspective of the interlocutor. These instructions when applied properly, allow an interlocutor to identify the object in question. The syntactic structure, i.e. the words and the grammatical relations of the utterance, encode which concepts and categories should be used and how the instructions work together. For instance, the fact that the hearer's perspective on the scene should be taken is conveyed by the phrase "von … aus" (from .. your perspective).

Languages vary widely in how they solve the problem of spatial reference – including both how they conceptualize space and how they talk about it (Levinson & Wilkins 2006; Levinson 2003). Spatial position of objects can be expressed using a variety of syntactic means including case, adpositions, particles, and verbs. But, maybe more importantly, there is a breathtaking variety in how people conceptualize space, which spatial relations they know, what counts as a landmark, how perspective is used, etc.. Just to give a few simple examples, Spanish has three basic proximal distinctions, while German has two. In Barcelona people make active use of the topology of the surrounding landscape, referring regularly to the seaside and mountainside when giving navigation instructions. In other languages *uphill* and *downhill* are used to refer to proximal objects (Levinson 2003).

Seite: 3

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:52:53
Langinfo missing

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:53:11
General Remark: Use small caps for abbreviations here

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:53:41
maybe small caps are better suited than console-style text

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:54:02
missing ,

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:54:55
There are two dots in the parantheses, but three in the preceeding quote.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:55:49
Here I think we may use only one .

These examples show that spatial language is a highly developed tool for establishing reference in a spatial environment. How did spatial language become this way? There is an emerging view now that the most plausible answer to this question is that spatial language is a *complex adaptive system* (see Steels 2000b for the general idea of language as a complex adaptive system), that is constructed and changed by its users for the same purpose it is used for today, namely to describe spatial scenes, establish reference to objects in the environment, give instructions for navigation, etc. This process is, of course, not the same process of construction that a group of engineers use when they are building a bridge. In such classic engineering problems, a team of people with a more or less complete view of the problem designs a top-down solution. By contrast, nobody has a global view on the state of a language. Rather, language lives in the individuals of the language community. Every individual has its own views on the state of the language, i.e. what words and grammatical relations are available.

When we combine the evidence from the complexity of particular spatial languages, such as German locative phrases, and the variation that can be seen across languages, it seems reasonable to consider results from a science that routinely deals with complexity and variation – biology. Biological species are highly complex solutions to particular environmental and social challenges. The solutions found by each species exhibit a high degree of variation. This simple observation has forced biology to come up with precise models and predictions to explain the origins of species. It comes as no surprise, then, that theories of language, particularly language evolution and language change, have adopted concepts from biology related to variation, complexity and the emergence of order in biological systems.

This book defends the *selectionist theory of language evolution*, which exploits biological concepts to explain how language is shaped by the communicative needs and environmental conditions that a community or population faces. The theory hypothesizes that agents create variation within their language and select working solutions based on how successful they are in communication (*communicative success*), how complex they are in processing (*cognitive effort*) and other factors.

Studying language change from the perspective of communicative intentions requires a great deal of insight into how humans or artificial systems can realize their specific communicative intentions in social interactions in the physical world. Such holistic explanations necessitate a *whole systems approach* (Steels 2001), in which great care is taken to ensure that perception, conceptualization and linguistic processing systems are integrated to an extent that interaction be-

Seite: 4

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 13:58:45
As this is an important term in its first occurence, the Generic Style Rules would recommend small caps

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:01:01
s.a. on small caps

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:00:57

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:00:53

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:00:48

tween agents is possible. Only when all of this machinery is in place can one attempt to examine questions of language change.

In particular, a whole systems approach requires an operational theory of language. How are utterances processed? How is space conceptualized? How is linguistic knowledge represented? How does language interact with the perception of the physical reality? A whole systems approach requires concrete answers to each of these important questions. The resulting burden placed on operational models is of course far greater than for high-level explanations or logical reasoning about these processes. But concrete, mechanistic accounts allow much greater insights into the phenomena studied. In the best case, a successful model of language evolution in a whole systems approach validates many aspects of the theory of language and language change at the same time.

This book contributes to the understanding of spatial language in two ways. First, it provides a detailed operational reconstruction of German locative phrases using a whole systems approach. Second, it explores the evolution of spatial language within the same computational framework. The two parts together argue for (1) the validity of the approach to language, and (2) the validity and explanatory power of the selectionist theory of language evolution.

1.1 Locative Spatial Language

If one wants to make an interesting claim about how language evolves, one needs a solid idea what language actually is, how linguistic knowledge is represented, and how to organize linguistic processing. These questions are best answered by reconstructing a complex natural language phenomenon such as German locative phrases. Such phrases are used for establishing reference to static objects and identifying them by denoting their spatial position (Miller & Johnson-Laird 1976). They can be distinguished from other parts of spatial language that are dealing with motion or navigation (Eschenbach 2004).

German locative phrases can be analyzed in terms of components or systems which together form a locative phrase. Example 1.1 consists of three parts: a *spatial relation*, which is combined with a *landmark* and a *perspective.*

Spatial Relations The defining quality of locative spatial phrases are that they contain locative spatial relations such as "rechts" (right), "vorne" (front), "nah" (near), "nördlich" (north on and so forth). These relations are called locative because they encode static spatial relationships and do not refer to change of position in time. In the Example phrase, "rechts" (right) is the

Seite: 5

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:04:30

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:04:40
maybe small caps are better suited here

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:13:56
General: According to the Generic Style Rules, the example should be set in italics and the translation in parantheses and single quotation marks

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:14:53
Do you mean to refer to E. (1.1) here?

locative spatial relation. In this book we study three classes of spatial relations. *Proximal* relations are based on distance estimations. Examples of proximal relations in German are “nah” (near) and “fern” (far). The second class is called *projective* relations and includes direction-based spatial relations such as “links” (left) and “vor” (front). The last class considered are absolute relations such as “nördlich” (north) and “östlich” (east). These are also direction-based, but the direction is related to a geocentric reference system such as the magnetic poles of the earth.

Landmarks A spatial relation is at least a binary and always relates to something. This something is typically called *landmark*. In ~~the~~ Example 1.1, the landmark is expressed in the determined noun phrase “der Kiste” (the box) immediately following the spatial relation.

Perspective For certain spatial relations perspective is important. Example 1.1 features a perspective that is marked via the phrase “von … aus ” (from … viewpoint). The marker expresses that the viewpoint on the scene is the hearer.

1.2 A Theory of Language Evolution

Theories of language evolution have to explain the evolution of language by defining the role and contribution of four different factors on language: biology, cognition, social cognition, and culture (Steels 2009, 2011c).

Biology To study language evolution from the biological perspective is to ask questions about the relationship of biology, in particular, genetics and ecology with linguistic behavior. The question can be roughly split into two parts. First, what is the biological influence on the general capacity for language in the human population? Second, one can ask for the influence of biology on the particular language spoken by individuals. The first is a general question for the processing capabilities that need to be present for language. This includes that humans require sufficient memory and powerful neural circuitry for processing language, But, also production organs for speech and auditory capacities. The second question is ~~one about~~ how much the biological basis determines the particular language individuals speak. In other words, how much the lexicon and/or the grammar of a language are influenced by genetic conditions.

Seite: 6

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:15:30
s.a. on italics and single quotation marks

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 14:15:56

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:17:10

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:17:08

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:19:09
I think it would be better to say that the factors are biological, not that biology (the science) itself was a factor of language.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:20:20
"are necessary for" ?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:21:18
capitalisation and punctuation do not match here. Do you want to have a . after "language" or a , and a regular "but"?

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 14:21:56

Cognition Biology has provided us with neural circuitry that enables distinct cognitive capabilities. The cognitive perspective on language asks: what are the basic cognitive processing mechanisms underlying production and parsing of language, interpretation, conceptualization, but also categorization, perception etc.? Language depends on a number of capabilities that may or may not be prior to language, such as temporal clustering of events, spatial navigation, perception-action systems (Rizzolatti & Arbib 1998; Arbib 2002; Steels & Spranger 2012, 2008), memory and so on and so forth. For instance, some have linked the evolution of language to an increase in capacity for storing cognitive categories and their interrelations (Schoenemann 1999). Another strand of cognitive influences on language evolution are general cognitive operators such as analogy, and learning operators, for instance, sequential learning (Christiansen et al. 2001).

Social Cognition Inevitably, language is a social phenomenon that occurs when humans interact. Social cognition researchers, for instance, are interested in the social mechanisms that are needed for children to acquire language, but also in the social mechanisms that are prerequisite for the emergence of language. Proposals include things such as *theory of mind* (Dunbar 1998) which is the capacity to understand another individual's state of mind, *joint attention* (Carpenter et al. 1998) which is the ability to track interlocutor gaze and mutual attentiveness to the same object, *social learning skills* such as imitation learning (Tomasello 1992) and the ability and the urge to *share intentions* (Tomasello et al. 2005). Many of these mechanisms are deeply rooted in biology. For instance, Dunbar (2003) and Worden (1998) argue that theory of mind is a necessary preadaptation for language and that it has evolved via natural selection.

Culture Language is a cultural phenomenon that is undergoing steady change on the cultural level. New words, speech sounds, morphemes, semantic and syntactic structures arise all the time in language (Steels 2011c). This manifests in the incredible amount of cross-cultural variation on all levels of linguistic processing (Evans & Levinson 2009), for example, phonemes (Maddieson 1984; Oudeyer 2005), spatial semantics (Levinson 2003), and syntax (Levinson & Wilkins 2006). This evidence points to strong cultural negotiation processes in which continuous invention is channeled to produce complex useful communication systems. Many of such processes orchestrating change and diversification have been identified. Grammaticalization, for instance, tries to explain the shift from lexical items to grammat-

Seite: 7

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:25:28
This appears ambigious to me. Maybe it would be better to write: "Nature has provided us..."
Confusion might arise since biology as a science helps us understand, but did not build our brain.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:28:29
I do not understand the meaning of "but" here. What is the difference between "perception" and "interpretation"? Is it because categorization may happen without a language basis, but interpretation is only possible "on top of a" language basis? Maybe it would work well to explain the difference between the first group (before "but, also..") and the second group (after "but also")

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:31:16
The punctuation in this recital is a little confusing.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:34:49
I think that it is unclear here whether you refer to biology as a discipline or to the biological properties of humans

Verfasser: fkopecky Thema: Notiz Datum: 28.05.2015 14:36:44
There are three hyphenations in a row here, which should be corrected

ical items (Hopper & Traugott 2003). Others have pointed to generational change as the trigger for development in language (Smith et al. 2003). The question from the perspective of cultural evolution is what are the mechanisms that bring about change in language and what are the principles with which agents conventionalize language up to the point that interlocutors have a chance of understanding each other.

I emphasize the cultural point of view in this book. That is, my primary concern is with change in language on the cultural level independent of changes in the human biology. Language change occurs on a smaller time scale than, for instance, the adaptation of a new ~~biological~~ organ, let alone a new species. There is absolutely no doubt that languages evolve fast. One just has to look through a text by Shakespeare or Goethe to see that a few hundred years can have impact on vocabulary and grammatical structure. It took Vulgar Latin a mere 1500 years to evolve into about a dozen different languages such as French, Italian , Portuguese or Catalan (e.g., see Pope 1952 for French). If we observe languages today, we can easily see that new words are invented all the time. In academic and technological contexts, for instance, new concepts arise all the time. Roughly 30 years ago vocabulary such as “email” or “website” did not even exist. What drives change in language, in what circumstances does it take place and what are necessary requirements for language change to occur? These are questions that cultural evolution theories of language have to address.

1.2.1 Language Systems and Language Strategies

Cultural theories of language evolution have to take a close look at individual trajectories of language change (Steels 2011c). For instance, how did the Russian aspectual system emerge or why does English have a system of determiners and Russian not? How do spatial language systems develop over time? In other words, cultural theories of language evolution must provide models for the emergence and evolution of concrete *language systems* (Steels 2011c). Language systems package a particular *semantic system* (e.g. a set of spatial categories) and a particular way of expressing these distinctions (e.g. a corresponding set of lexical items). The absolute German system, for instance, consists of four absolute spatial categories and the corresponding strings, e.g. “nördlich” (north), “südlich” (south). These spatial categories are the basic building blocks of absolute spatial conceptualization in German. They can be compositionally combined with landmarks to build complex spatial phrases. Interestingly, the German locative systems, effectively consists of different conceptualization strategies that have

Seite: 8

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 14:38:46

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:38:36
maybe "development" is a better term here

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:40:33
extra spaced

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:42:17
Section 1.2 is not exhaustive - it is not the case that all text in section 1.2 is aligned to a subsection.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:43:24
see above on markup for in-line examples and their translation

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:43:56
I think this , should be omitted for better reading flow

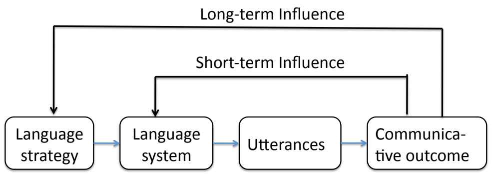

Figure 1.1: The fitness of utterances for communication affects both the language system and the language strategy. The effect of the success of a single utterance on the language strategy is smaller which leads to slower change on the level of the strategy. (Figure adapted from Steels 2011c)

distinct but converging evolutionary trajectories. For instance, the absolute system is connected to the invention of the compass, whereas projective systems often at least in part can be traced back to body parts (Traugott & Heine 1991). Nevertheless, many locative spatial relations are used in the same syntactic context.

Spatial language systems such as the proximal or projective system are characterized by a degree of cohesion and systematicity that points to an underlying principle that organizes acquisition, emergence and coordination. We call the mechanisms organizing a particular language system the *language strategy* (Steels 2011c). Language strategies have a *linguistic* ~~part~~ and a *conceptual* part. For example, on the conceptual side absolute spatial categories share that they are part of the same conceptualization strategy which uses absolute directions to the magnetic poles of the earth. Syntactically all spatial relations share that they are expressed in a similar way namely lexically and that they can be expressed as adjective, adverb and preposition.

1.2.2 Selectionist Theory of Language Evolution

In this book I follow the *selectionist theory of language evolution* (Steels 2011c), which applies the dominant theoretical construct in biology *Natural Selection* and uses it to explain language change. on the level of *language systems* and *language strategies*. Additionally, the concepts of *self-organization*, *recruitment*

Seite: 9

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:46:35
Maybe small caps would be better suited here?

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 14:46:10

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:48:03
The capitalization appears incoherent to me considering the non-capitalized terms from linguistics.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:49:00
The . here should be deleted in order for the next sentence to make sense.

and *co-evolution* of syntax and semantics are used as theoretical pillars.

Selection Selectionism rests on two principles: *generation* of possible variants and *selection* of variation based on fitness. The most important factor in determining the fitness of a particular language strategy, but also of a particular language system, is *communicative success*. A communicative interaction between two interlocutors is successful if the communicative intention of the speaker is reached. For instance, if the speaker wanted to draw attention to some object, the communication is successful if the hearer pays attention to that object. Communicative success drives selection on the levels of the language system, but also on the level of language strategies (see Figure 1.1).

Variation occurs in the systems for two reasons. First, agents are actively trying to solve problems in communication (Steels 2000b). Agents introduce new categories, new words and grammar when they detect problems that they cannot solve using the current language they know. Second, language is an inferential communication system (Sperber & Wilson 1986) which means that the information provided in an utterance is often incomplete and ambiguous. Interpreting phrases is an active process in which the hearer is fusing information from the context, from the dialogue and his knowledge about the language to arrive at the best possible interpretation. In this process of course hearers might interpret the utterance differently then intended. This is the second source of variation.

Self-organization Steels (2011c) assumes that selection is not enough to explain language change and proposes another driving force in the evolution of language: self-organization – a concept used to account for complex phenomena in physical and biological systems. In short, self-organization is a way to explain how global structure arises out of local interaction of subunits (Camazine et al. 2003). An example from biology for self-organization is swarm behavior in school of fish. Each individual fish locally controls its behavior based on the estimation of the position and direction of its immediate neighbors. On the global level this leads to consistent swarm behavior. Self-organization is typically seen as a complementary mechanism to selection. Although there is some discussion on how to reconcile the two mechanisms. Kauffman (1993), for instance, proposes the following idea. Local components and the interaction rules are determined by selection, whereas the global emergent behavior is explained using self-organization. Applied to the swarm behavior this means that

Seite: 10

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:49:18
This line should be moved to the preceeding page.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:52:48
"in a school of" or "in schools of"

Verfasser: fkopecky Thema: Notiz Datum: 28.05.2015 14:55:08
Too many hyphened lines in a row

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:55:43
selection, although ??

the anatomy of fish as well as the perceptual feedback loop are a product of Natural Selection. The global emergent swarm behavior is the product of self-organization.

Similar to swarm behavior, agents in a population evolving a language have to achieve global coherence in the language they use. Each agent has its own private representations of the language that he speaks and he can adjust his own representations based on local interactions with peers. How, from local interactions, agents can agree on a globally shared communication system is the problem of alignment. Psychologists have found that interlocutors align on all levels of linguistic processing even over the course of a few interactions, i.e., dialogue (Garrod & Doherty 1994; Pickering & Garrod 2004). Similar mechanisms applied over a long time span are required for driving populations to self-organize a sufficiently shared communication system (Steels & Kaplan 2002).

Recruitment The last problem for an account of how languages change in the *selectionist theory of language evolution* is the problem of language strategy generation. The hypothesis is that language strategies are recruited by assembling basic cognitive operations (Steels 2007). For instance, an absolute spatial conceptualization strategy involving distinctions such as “north” and “south” consists of basic categorization mechanisms and the ability to track ones own direction. The two abilities are assembled into the strategy which encompasses the different absolute spatial distinctions. The process is called *recruitment* because the cognitive mechanisms which are assembled could, in principle, have evolved or could be learned independently from language.

Co-evolution One of the tenants of the theory of linguistic selection is that syntax and semantics co-evolve. The idea is that recruitment of conceptualization strategies and the invention of new semantic distinctions and spatial relations trigger evolution of the syntax of a language (Steels 1997, 1998). For instance, presumably when the absolute system in German emerged based on a new way of construing reality, this at the same time triggered the invention of new words.

1.2.3 Evolutionary Explanations

In every science one has to define what counts as an explanation. This book is guided by what counts as an evolutionary explanation in biology, ethology

Seite: 11

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:56:10
rather "natural selection"?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:56:35
maybe better: "they speak"

and psychology(Tinbergen 1963; Dunbar 1998). In order to explain a complex trait from the evolutionary perspective one has to provide explanations on four different levels: *function*, *mechanism*, *ontogeny* and *phylogeny*.

Function An explanation for a particular behavior has to show what the behavior is good for, i.e. what is its purpose. For Darwinian biology, the function of a behavior has to be explained in terms of its impact on survival or, more precisely, on the production of offspring. For evolutionary linguistics this turns into the question of how a particular language system or a particular strategy helps an agent to be more successful in communication. For example, one can explain particular spatial language systems with respect to their ability to help agents solve communicative problems in spatial navigation and spatial reference.

Mechanism Besides function, one has to identify the mechanisms that give rise to the behavior. This is actually called *causation* by Tinbergen (1963) and it refers to the cause and effect relations that generate a particular behavior. For instance, one can explain how aggressive behavior is generated by looking at changes in hormone levels in an organism, e.g., testosteron causes aggressive behavior. For spatial language this entails a detailed operational model of the production and parsing of spatial language.

Ontogeny The next question is how a particular behavior is acquired. To answer this question one has to identify the developmental steps that the behavior undergoes, but also what is the ontogenetic bases of the behavior. What is learned and what is instinct? For spatial language this requires insights into how spatial language is learned.

Phylogeny A fourth part of every evolutionary explanation has to identify the evolutionary history of a behavior. What are the sequential stages of evolution of a behavior? What are the prerequisites of a behavior? How do evolutionary older behaviors influence the behavior under question. These questions have to be answered with respect to the function of the behavior. In other words, one needs explanations of how the behavior evolved to fulfill its current function. For language evolution scholars have to identify how a particular strategy evolved over time. Was it adapted from an older strategy? How did syntax and semantics the language system under consideration co-evolve over time?

Seite: 12

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 14:58:55
There needs to be a space here.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:05:14
Wouldn't it be better to say ... "has to show what its purpose is" .. ??

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:07:39
either: "what are the ontog. bases" or "what is the ontog. basis"

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:08:27
This is a question, so it should be ended with a ?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:09:25
This sentence somehow says the same as the preceeding one. If this is so, it should be deleted.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:10:05
"syntax and semantics OF the language system" ?

1.3 Main Hypothesis

This book provides experimental evidence for the theory of linguistic evolution. The hypothesis is that *spatial language syntax and spatial semantics co-evolve through a cultural process based on selection, self-organization and recruitment*. This book explores the hypothesis for the different components of spatial language: spatial relations, landmarks and perspective. Computational experiments show the emergence of spatial relations, the negotiation of the use of landmarks and perspective. I also explore different strategies for expressing spatial conceptualizations: lexical and grammatical strategies.

1.4 Contributions

This book provides detailed accounts of the function, mechanisms, ontogeny and phylogeny of spatial language.

1. The first contribution is an explanation of the mechanisms behind spatial language for German locative phrases in a complete reconstruction including perception, semantic and syntactic processing. Once the mechanisms are in place, we test the function and impact of components of spatial language in experiments by removing the component in question and examining the effect the removal has on communication.

2. The second contribution is to explain steps in the co-evolution of spatial syntax and spatial semantics through computational models.

This book follows a whole systems approach which allows us to define external criteria for the progress in each of the objectives. The defining moment for the underlying conception of language is communication. A communication system is *successful* if it allows robotic agents to achieve their communicative goals such as drawing the attention to an object in the environment.

1.4.1 Evolutionary Stages

One way of understanding evolutionary processes is to try and identify evolutionary stages. Over the years, different steps in the evolution of language have been identified involving varying degrees of specificity (Bickerton 1999; Jackendoff 1999; Steels 2005). All of these proposals, while differing in detail and the exact number of stages, agree that language evolution starts at some pre-grammatical stage and increases in complexity to the form of language, in particular, grammar

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:16:03
to try to identify ?

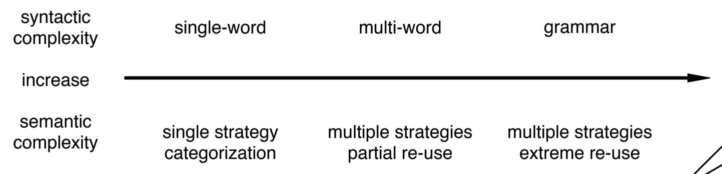

Figure 1.2: Co-evolution of syntactic and semantic complexity.

that we see today. Obviously any evolutionary account of language has to show how the current state of complexity of language can be traced back to earlier simple stages.

This book orients itself alongside Steels (2005) which proposes a number of stages of complexity which are relevant for this book: *single-word utterances*, *multi-word utterances*, *grammatical utterances* (see ~~also~~ Figure 1.2).

Single-word utterances In this stage agents utter single words that pertain to a particular concept or category used for discriminating objects. Examples for spatial language include utterances that directly refer to spatial regions such as “links” (left) or “nördlich” (north). When agents can only express themselves using a single word, this single word necessarily encodes the complete conceptualization strategy. For instance, which landmark or perspective is used for conceptualization is holistically coded in the single word. Since there is no additional information about which conceptualization strategy the term is referring to, agents have to implicitly agree on the precise spatial construal the term is referring to. This limits the re-use of spatial categories in different spatial conceptualization strategies because agents have no way of disambiguating the use of the same spatial relation in different strategies.

Multi-word utterances Single-word communication systems are not very flexible. There is no compositionality and particularly there is no re-use. In German, for instance, projective relations can be used with different landmarks. In the multi-word utterance stage agents can express different constituents by using a number of lexical items. Besides expressing the spatial category used, agents can also mark landmarks. An example utterance is “links

Seite: 14

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:17:32
rather "who"? The parantheses suggest that you are referring to the person Steels and not to one of his works.

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 15:17:57

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:17:51
, and grammatical ... ?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:18:15

Kiste” (left box) which is used to signal that the region left of the box is meant.

Grammatical utterances When we look at natural language, we can see that the *same* constituents can be part of different conceptualization strategies. Imagine an utterance like “Kiste link” (box left) without the grammatical information, in particular, without word order and lexical class information. In that case a hearer does not know whether “link” is an adjective or an adverb. This syntactic underdetermination has consequences for the semantic interpretation. If the phrase is interpreted as an adjective noun phrase as in “linke Kiste” (left box), the spatial category acts as a modifier on the set of boxes. If the spatial relation is interpreted as an adverb, then box might be a landmark and the whole phrase denotes a region next to the landmark as in “links der Kiste” (to the left of the box). Grammar signals the difference in these two semantic interpretations and disambiguates the conceptualization strategies. Consequently, agents equipped with grammatical strategies can disambiguate even more strategies and consequently, they can be more expressive.

The goal of this book is to identify, implement and test the mechanisms that drive the evolution of language on *each* of these stages. The mechanisms we are interested in are not descriptions of the phenomena but mechanistic explanations which identify the computational and cognitive components that enable robotic agents to self-organize communication systems. The procedure to find and validate mechanistic explanations is to

1. hypothesize *invention*, *adoption* and *alignment* operators for the syntax and semantics according to each stage of complexity,
2. equip agents with these operators,
3. test the evolutionary dynamics in populations of equipped agents,
4. measure the communicative success, adaptivity and expressivity.

Invention, adoption and alignment operators are the backbone of the evolutionary models of this book. For instantiating the theory of linguistic selection one has to identify agent-level mechanisms that orchestrate the global behavior of the population. The mechanisms can be classified into the following three classes.

Invention operators Invention is the process of introducing variation into the system by inventing a new spatial relation or a word or even grammar in order

Seite: 15

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:19:30

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:19:34

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:20:16
rather "box" (in ") ?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:20:21

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:21:25
should Punctuation be used in a numbered list?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:21:10

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:21:13

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:21:15

Number of conceptualization strategies A second notion of semantic complexity is the number of conceptualization strategies a language supports. German, for instance, supports many different categorization systems: projective, e.g. "links" (left) or "rechts" (right), proximal, e.g. "nah" (near) and "fern" (far), and absolute, e.g. "nördlich (north) and "südlich" (south). This is one aspect. The other aspect is that these systems are part of different conceptualization strategies. Examples of this re-use were already given earlier with respect to adjectives and adverbs.

1.5 Structure of the Book

This book is structured into three main ~~big~~ parts besides this introduction and the conclusion. The first main part explains the interaction model and the technical systems needed for studying spatial language. The second part deals with objective number one and details the reconstruction efforts for the German locative system. In the third part I detail how spatial language evolves based on the model of evolutionary stages.

1.5.1 Part I: Spatial Language Games and Technical Background

Spatial Language Games Spatial language occurs mainly in interactions of individuals in spatial scenes. To research spatial language in such a communication-based approach to language a number of things need to be in place. We need a model of interactions in spatial scenes. This is the topic of Chapter 2 which introduces *spatial language games* which are routinized interactions consisting of defined roles for interlocutors – speaker and hearer. The chapter explains the basic interaction scheme and the linguistic and non-linguistic behaviors that define a spatial language game.

Embodied cognitive semantics with IRL In order to achieve the objectives of this book, we need computational formalisms that support the reconstruction and evolution investigations. One of such formalisms in part developed for this book is the *Incremental Recruitment Language* (IRL). IRL is (a) a formalism for representing semantics, (b) a set of planning algorithms for automatic conceptualization and interpretation, and (c) a set of tools that make semantics an open-ended adaptive system. Chapter 3 introduces the formalism and the technology behind it.

Seite: 17

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:23:35

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 15:23:55

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:26:26
maybe reference this part.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:24:25
"the first objective"?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:27:30
here, \beign{description} is not indented which it is in other occurences!

Construction grammar with FCG Another important backbone of the investigations is Fluid Construction Grammar (FCG). FCG is a formalism for representing and processing linguistic knowledge. Chapter 4 details how mappings from semantics to syntax are implemented using FCG and gives an example of processing a simple phrase.

1.5.2 Part II: Reconstructing German Locative Phrases

To ground the modelling efforts in sufficient knowledge of a real spatial language system, I decided to reconstruct a part of German spatial language – German locative phrases. The second part of this book reconstructs the syntax and semantics of German locative phrases. The part starts out with an in-depth look at German locative spatial language as a natural language phenomenon. Chapter 5 gives more examples of the syntactic variety and the connection to the space of conceptualization strategies supported in German locative phrases. This sets the scope for the reconstruction effort, but also identifies a number of processing issues that the reconstruction has to deal with in order to be successful.

Spatial semantics The following chapter details the operationalization of spatial semantics. Chapter 6 the basic semantic building blocks of German locative phrases and discusses how they work together to make up the complex semantics of spatial scenes.

Syntactic processing A close look at German locative phrases reveals a number of interesting phenomena. Most importantly it uncovers the tight relationship between spatial syntax and spatial semantics. Chapter 7 explains how FCG can be used to model the tight connection between the words and grammatical relations observed in German locative phrases and the world of spatial semantics. These mappings are interesting because they pose particular challenges to the organization of linguistic processing. The re-use of the same spatial categories in different strategies for conceptualizing reality and their syntactic expression requires sophisticated mechanisms for dealing with many-to-many mappings in language. Another important issue is how to deal with the case system of German. All of these aspects of linguistic processing are discussed in Chapter 7.

Conceptualization of spatial scenes Spatial scenes do not come a priori labeled, categorized and construed. Agents have to autonomously conceptualize reality given the particular communicative goal they have. Chapter 8 deals with

Seite: 18

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:28:07
modelling is en-UK, modeling would be en-US

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:29:12
There is a verb missing here.

Language strategies are tested in two scenarios – *acquisition* and *formation*. In acquisition a learner agent has to pick up the spatial language system spoken by a tutor. In formation all agents start from scratch and progressively develop categories and lexical items.

The most important influence on what kind of language system emerges is the language strategy. The chapter details different language strategies necessary for building proximal, projective and absolute systems which encompass dedicated invention, adoption and alignment operators as well as the different conceptualization strategies. The success of the learning operators and the conceptualization strategy is tested in experiments where populations are fitted with a particular strategy. The resulting languages spoken by individual agents are analyzed with respect to communicative success and how similar they are to each other.

Another important factor influencing the emerging language system are environmental conditions. The chapter studies the impact of environmental conditions systematically by manipulating environmental features such as global landmarks or the statistical distribution of objects.

Obviously, natural languages support many conceptualization strategies at the same time. German, for instance, simultaneously has a proximal, a projective and an absolute system. So one can ask what happens when agents are simultaneously operating different strategies. I hypothesize that agents need additional cognitive mechanisms for choosing between different strategies and that choosing a strategy can be realized using the *discriminative power* of each strategy in a particular context. When an agent has to invent a new category he uses the strategy that is most discriminating using a new category. Experiments show that this principle allows agents to build multiple language systems at the same time. Lastly, the chapter also studies the impact of different environmental layouts on formation of language systems for interacting strategies.

Origins and alignment of spatial conceptualization strategies Chapter 11 deals with the emergence and alignment of conceptualization strategies When one compares different languages of the world it becomes clear that many languages differ in the kinds of conceptualization strategies they support. Some languages solely use an absolute system, others can use intrinsic and relative systems and so on and so forth. Consequently, the evolution of spatial language is intricately connected to the origins and evolution of spatial conceptualization strategies. The chapter shows that conceptualization strategies are organized in a process of recruitment, selection and self-organization.

Seite: 20

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:35:13
maybe better: "they use"

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:35:32
missing .

To explain conceptualization strategies from the viewpoint of the theory of linguistic selection is to explain (a) how different conceptualization strategies are created and (b) how they are selected for in communication. Competition is an important aspect of selection. Obviously environmental conditions and communicative success are main influences on which strategies are selected for because they are more successful. The chapter proposes alignment operations that update and track the score of conceptualization strategies so that agents can locally align in their interactions. I show that these operators lead to global convergence of the population on using single conceptualization strategies. The chapter studies competition of different strategies for landmarks and frames of reference and shows that with the right alignment strategy agents can agree on using a particular conceptualization strategy while co-evolving a lexicon and ontology of spatial relations at the same time.

Besides selection the theory has to explain how conceptualization strategies are created. This is were the idea of recruitment comes into play. Conceptualization strategies are assemblies of cognitive operations. For instance, an absolute strategy consists of a particular way of applying spatial categories plus the computation of a global landmark. Recruitment is the process of drawing from the pool of cognitive operations and assembling and packaging them so that the complete structure for conceptualization can be scored and the score updated and tracked. In a second set of experiments creation and competition of strategies are studied together.

Multi-word lexical systems for expressing landmarks Single-word utterance systems are limited in how much information can be conveyed in them. Upon hearing a single term it is hard to decide what conceptualization strategy was it part of. Which landmark is used? Which perspective did the speaker have in mind? These are questions that cannot be decided by just looking at a single word, unless of course the word is known and always refers to the same landmark and the same conceptualization strategy. When we look at human language we see a lot of re-use of spatial relations. Absolute, projective and proximal relations in German can be used with different landmark objects. Chapter 12 examines what mechanisms are needed for agents to mark landmark objects using lexical items while at the same time co-evolving a lexicon and ontology of spatial relations. Once these mechanisms are in place success of such extended lexical systems can be studied and compared to systems which only support a single conceptualization strategy.

Seite: 21

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:36:53
I think one of "for" and "because" is superfluous

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:38:05
Make these into subsubsections?

Grammar as a tool for disambiguating spatial phrases The part on language evolution of this book is concluded by a chapter that examines the role and evolution of grammatical language.

Lexical systems which are all systems studied up to this point in the book, have considerable shortcomings. One can study the effect grammar has by removing grammatical knowledge from the German locative grammar implemented for this book. The results presented in Chapter 13 show that agents operating a German locative system without grammar have significantly lower communicative success. I show that environmental conditions and diverging perspective on the scene can increase the drop in communicative success. The lack of grammar increases *semantic ambiguity* of phrases which means that the number of possible interpretations of a phrase escalates. As a consequence, the number of wrongly interpreted topics enlarges as well.

Given such a clear communicative advantage for having grammar, one can study the necessary operators that enable agents to develop a grammar for disambiguating spatial phrases. This is the topic of the second part of Chapter 13 which reports on the precise implementation of these operators. I test the operators in multi-agent experiments which prove that the hypothesized invention, learning and alignment operators allow agents to become increasingly more successful in communication because they develop an effective grammatical communication system.

Seite: 22

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:38:48
which one?

2 Grounded Spatial Language Games

Language does not occur in a vacuum. Spoken language occurs in physical, situated interactions when two interlocutors meet with specific communicative intentions. This chapter explains the basic social interactions at the center of the approach to language. Physical robots meet in communicative encounters and try to reach communicative goals within real world settings. Taking such a radical approach to the study of language is grounded in a number of social and perceptual mechanisms. Here I look at the prerequisites for the computational models discussed later in this book.

Figure 2.1 shows such an encounter of two humanoid robots in which one of the robots, the speaker, has the goal of drawing the attention of the hearer to some object in the environment using language. Such interactions are called *language game* (Steels 2001). Language games are routinized interactions between two members of a population. The game combines a particular script for the interaction, the linguistic information transfer and extra-linguistic feedback about the success of the interaction. Here is an example of a language game called the *spatial language game*.

1. The language game starts by randomly drawing two agents from the population. One agent is randomly assigned the role of speaker, the other is assigned the role hearer.
2. The agents establish joint attention and perceive the scene.
3. The speaker chooses an object from the perceived context as the object he wants to draw the attention of the hearer to.
4. The speaker produces an utterance that he thinks draws the attention to the object.
5. The speaker passes the utterance to the hearer.
6. The hearer interprets the utterance and tries to find the object that the speaker might have in mind.

Seite: 25

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:49:05
The sectioning of this chapter is not exhaustive.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:49:59
"games"?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:49:53

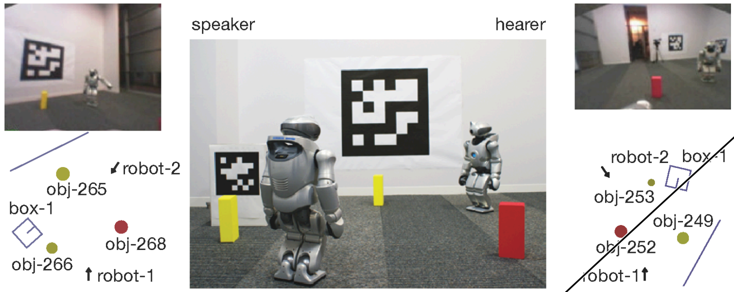

Figure 2.1: Example scene. Two robots autonomously perceive and act in an office environment that contains different types of objects. Both robots autonomously create world models reflecting the state of the environment (see bottom left and right schematics), that include objects with spatial and color properties, the carton boxes as well as the robots.

7. The hearer points to the object he thinks the interaction was about. If he was unable to interpret a topic, he signals this by shaking his head.

8. The speaker interprets the pointing. If the object pointed to by the hearer is correct, he signals this by noding. If the hearer pointed to the wrong object or did not point at all, the speaker points to the topic he wanted to draw the attention to.

To study language in a real world setting requires to fully spell out all components involved in the interaction. Besides social mechanisms agents need operational systems for perceiving the world, as well as the construction and interpretation of utterances. Figure 2 shows a schematic view of the systems involved in production and parsing, conceptualization and interpretation. Both agents independently process sensorimotor data stemming from the onboard cameras and proprioceptive sensors in order to construct world models of the environment (Spranger 2008; Spranger et al. 2012a). Based on the particular communicative goal and the current state of the world represented in the world model, the speaker conceptualizes a meaning which is then rendered into an utterance by

Seite: 26

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:51:25

Agents perceive these objects as having a certain distance from their egocentric coordinate systems originating in the robot's body.

Boxes The environment also features card boxes which have particular markers on their sides. These markers are perceived by the sensorimotor systems as distinct sides of the object. The perceptual systems perceive them as having a particular distance and orientation with respect to the robot's coordinate system.

Wall markers Figure 2.1 shows that the same markers used for the carton boxes can also occur on the wall. Cardboard boxes introduce a geocentric orientation on the scene.

Robots Robots also establish the position of the interlocutor in each interaction. Every robot tracks the position and orientation of the other robot in his environment.

Before a language game starts the robots perceive their environment. The robots are endowed with perceptual systems for recognizing and tracking the objects in their environment. These systems continuously build up *world models* of the environment consisting of sets of objects. The objects are characterized by continuous real-valued features such as color, position and orientation but also width, height and length. The perceptual system also provides a basic grouping of objects into classes such as robots, blocks and boxes and wall markers. The following is the world model built by the agent to the left in Figure 2.1. It includes the other robot (`robot-2`), the landmark (`box-1`) and the colored blocks (`obj-265`, `obj-266` and `obj-268`):

```
((robot-1 :type robot :x 0.0 :y 0.0 :orientation 0.5)
 (robot-2 :type robot :x 1461.65 :y -351.24 :orientation 0.9)
 (box-1 :type box :x 513.0 :y 891.67 :orientation 0.36
        :width 320.0 :height 450.0 :length 310.0)
 (obj-265 :x 1454.74 :y 248.72 :z 0.0 :width 59.75
          :height 235.99
          :average-y 128.0 :stdev-y 26.49 :min-y 51.0
          :max-y 199.0 :average-u ...)
 (obj-266 :x 285.0 :y 549.02 :z 0.0 ...)
 (obj-268 ...)
 (box-wall :orientation 0.36))
```

Each robot constructs perceptual representations of the objects in its immediate surroundings from the raw sensations streaming from the robot's sensor. Each type of object in the environment is tracked by a dedicated perceptual system. In general, processing of the different object classes is a three step process. First, low-level vision routines process raw camera images to yield basic *percepts*

Seite: 28

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:56:48
Line spacing has been changed in this par. Has this happened by accident?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:56:53
,

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 15:57:38
maybe these categories should be set in small caps for better readability

2.2 Social Mechanisms

Language games require a number of social mechanisms to be in place. Joint attention, turn-taking behavior, pointing and other non-linguistic feedback ~~these~~ are mechanisms at the heart of the social interaction. Crucial social mechanisms required for these interaction are considered prerequisites for studying the evolution of language. They constitute the background against which communication and evolution of communication takes place.

Joint attentional scenes (Tomasello 1995) in which interlocutors are jointly attending to some object for some reasonable amount of time. Establishing joint attention means in robotic experiments that two robots taking part in a language game must (1) share a physical environment, (2) attend to a set of objects in their surrounding, (3) track whether the respective other robot is able to attend to the same set of objects and (4) be able to manipulate attention by pointing to distal objects and perceiving these pointing gestures Joint attention is monitored by an external computer program, that has access to the world models of both interacting robots. This system initiates the interaction between two agents as soon as both agents observe the same set of objects. Spatial scenes are manipulated by a human experimenter to find spatial setups in which joint attention is possible, the program monitors whether robots are seeing the same set of objects and informs the experimenter whether the robots jointly attend to the same set of objects.

Social interactions have to be structured, so that agents can interpret the signals they exchange. For instance, if the hearer points before he has received the utterance the speaker will have a hard time understanding the gesture. However, if the hearer points after receiving the utterance, the speaker can assume that this is the response to his speech act. Language games are coordinated by behavioral scripts. Every agent in the population knows the language game script and individually reacts to changes in the environment and actions of the other robot. For example the speaker triggers the action of pointing to the intended topic when the hearer signals that he did not understand the utterance. The scripts are implemented in the form of finite-state machines: actions are performed depending on the current state in the game flow, the perception of the environment and the history of the interaction.

In order to be able to learn and form spatial language systems, robots need non-linguistic means of conveying information, such as pointing to an object or conveying notions of success, failure and agreement in communication. For demonstration purposes robots are equipped with pointing gestures but in the

Seite: 32

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:03:44
better: to work ?

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 16:04:22

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:05:14
This sentence is missing something.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:05:48
For better readability, I'd like to suggest moving this to the beginnging of the sentence

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:06:20
There is a . missing here

communicative interactions underlying the results presented in this book, robots use a different mechanism in order to avoid further difficulties stemming from uncertainties in pointing (see Steels & Kaplan 1998 for a discussion of the impact of such uncertainties on the performance in language games). When a robot wants to point to an object in the environment, he directly transmits the coordinates of the intended object to the interlocutor. Since robots model object positions in their own (egocentric) coordinate systems, additional steps have to be taken to interpret these coordinates. Most importantly the robot has to know the position and orientation of the robot that is pointing. With this information robots transform the coordinates into their own coordinate system and interpret the pointing by choosing the closest object to the pointing coordinates in their world model. Similarly, robots directly exchange other non-linguistic feedback, for instance agreement and disagreement in communication by exchanging signals whose meaning is shared. Moreover, linguistic utterances are directly passed between interlocutors.

Seite: 33

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:09:09
"they" or "it"?

3 Embodied Cognitive Semantics with IRL

Artificial agents trying to achieve communicative goals in situated interactions in the real-world need powerful computational systems for conceptualizing their environment. In order to provide embodied artificial systems with rich semantics reminiscent of human language complexity, agents need mechanisms for both conceptualizing complex compositional semantic structure, but, also for actively reconstructing semantic structure in interpretation of ambiguous utterances. Furthermore, the system must be open-ended and allow agents to adjust their semantic inventories in order to reach their goals. This chapter presents the computational system called Incremental Recruitment Language (IRL) that allows agents to represent and process complex conceptualizations of spatial scenes. The work presented here bases itself on substantial previous work. Key ideas of the IRL system have been laid out by Steels (2000a), with progress reported in Steels & Bleys (2005), Van den Broeck (2008), and recently in Spranger et al. (2010a, 2012b).

3.1 Procedural Semantics

In order for a hearer to interpret an utterance, he has to apply the meaning conveyed in the linguistic structure to his perception of the context. Consequently, a speaker who uses language to achieve a certain communicative goal wants the hearer to execute a program (Johnson-Laird 1977), i.e. a set of operations that allow the hearer to, for example, discriminate an object in the environment or perform an action. Thus we model semantics, i.e. what it is a speaker wants the hearer to execute, as a program linking operations and data. Let us start with an example. Suppose a speaker utters the phrase "der rote Block" (the red block) with the intention of making the hearer point to an object. In this case, the phrase encodes a program, i.e., set of operations, that are supposed to lead the hearer to identify the object in question. Presumably the hearer of this utterance has to first filter the context for blocks, followed by the application of the

Seite: 35

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:12:02
real world

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:12:59
The combination of "both" and "but" might produce misunderstandings.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:12:39

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:16:47
see previous comments on example layout

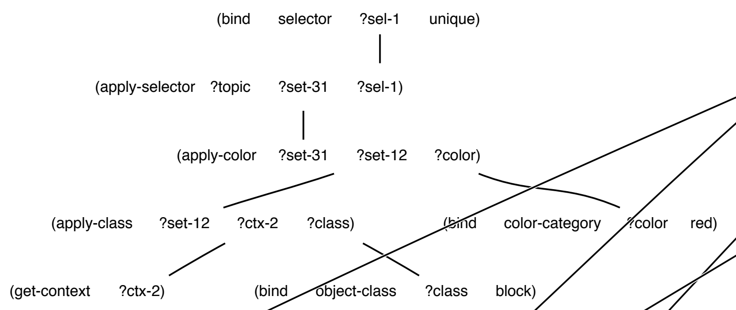

Figure 3.1: Semantic structure underlying the utterance "der rote Block" (the red block).

color category red, in order to arrive at the set of red blocks, which is used to compute the topic consisting of a single entity. A possible program, also called *IRL-network*, is shown in Figure 3.1. This network explicitly represents the chain of the four operations `get-context`, `apply-class`, `apply-color` and `apply-selector` by linking their arguments through variables (starting with `?`). The network also includes the color category `red`, the object class `block` and the selector `unique` which are introduced via so called *bind statements*, as in `(bind color-category ?color red)`. We collectively refer to concepts, categories etc. as *semantic entities*.

IRL-networks consist of two types of nodes

Cognitive operations, also called *semantic operations*, are the algorithms used in conceptualization. They encode a particular cognitive function such as categorization using a color category, applying a selector or applying an object class and many more as will be shown see later in this book for the domain of space. Cognitive operations are identified by their name, e.g. `apply-color` and they have a set of arguments which can be linked to other operations or semantic entities via variables (starting with `?`).

Semantic entites is the general term for referring to prototypes, concepts and categories that are used by cognitive operations. Besides such long-term

Seite: 36

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:17:05

Verfasser: fkopecky Thema: Notiz Datum: 28.05.2015 16:19:02
The distance between Figures and the Text appears unusally wide to me.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:17:41
maybe small caps would be better suited here for readability?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:17:23

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:17:25

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:17:58
"of nodes:" ?

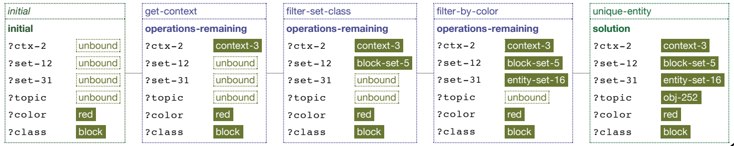

Figure 3.2: Progressive evaluation of the network in Figure 3.1 on the context shown in Figure 2.1. From left to right, each node represents a step in the evaluation process. From top to bottom, the evaluated operation, the node status, and the current list of bindings of each node are shown. A consistent solution with bindings for all variables is found in the last node, and the value `obj-252` is indeed a unique red block (compare 2.1).

data, semantic entities can also be discourse representations, the representation of the current context and data exchanged between cognitive operations. They are introduced explicitly in the network via `bind`-statements which are special operations for retrieving the actual data representation using a pointer or shorthand notation for it. For instance, the statement `(bind color-category ?color red)` encodes the access to the color category `red` which is a prototype represented using values for different color channels.

3.2 Evaluation

A program such as the one in Figure 3.1 is *evaluated*, by a speaker to test the semantic structure with respect to the particular communicative goal, or by a hearer in order to interpret an utterance. Evaluation is a process which cycles over the network and progressively computes values for variables, a process called *binding*. When the network in Figure 3.1 is evaluated the following happens. First `get-context` gets the current world model from the perceptual processes that are monitoring the environment for events and objects and binds it to the variable `?ctx-2`. This is followed by the evaluation of the `apply-class` operation which computes a similarity score for how similar every object in the context is with respect to the object class `block`. This yields the set of objects

Seite: 37

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:19:57

from the context with each object scored using the computed similarity. The set is bound to the variable `?set-12`. Because this variable is linked to the operation `apply-color`, the set bound to the variable `?set-12` is further processed using the color category `red`. `apply-color` first computes a similarity score for every object in the input set to the color category `red` which is multiplied with the similarity score the object already has from the application of the class `block`. This yields a new set of objects with multiplied similarity scores. The set is bound to the variable `?set-31`. Lastly, `apply-selector` checks the objects in `?set-31`, finds the object with the highest similarity score and binds it to the variable `?topic` which is the referent[1] of the phrase "der rote Block" (the red block). Figure 3.2 gives an idea how variables get progressively bound when the IRL-network is evaluated.

This is only one example how such a network can be evaluated. As has been argued in Steels (2000a), language requires that semantic structure does not encode control flow, but rather data flows in all directions and is computed wherever possible. For this, operations need to be able to function in different directions with varying input-output parameters For instance, the operation `apply-class` which has three arguments, applies a class such as `block` to an input set, when the class is explicitly represented in the network. But, in case this class is not introduced via a bind statement in the network, the operation can also provide this information effectively turning this argument into an output argument. This *multidirectionality* of operations proves important for dealing with missing items, for instance due to partial parsing of an utterance, but it is also needed when constructing semantic structure.

3.3 Conceptualization and Interpretation

There are two scenarios in which agents autonomously compose semantic structure like the one just described (Figure 3.1). First, speakers have a particular communicative goal and need to construct semantic structure, for instance, for singling out the particular topic the speaker wants to draw attention to. This process is called *conceptualization*. In the second scenario, hearers use information parsed from the observed utterance and their knowledge about the current context of the interaction to actively reconstruct meanings from the potentially partial structures parsed by the language system. We call this process *interpretation*. Both cases are equally important and they both conceive the process of

[1] Note that the word "object" here refers to an agent's private representation of things he has perceived in the world and only indirectly refers to the physical object which is the referent.

Seite: 38

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:21:35

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:21:43

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:21:47

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:21:56

Verfasser: fkopecky Thema: Notiz Datum: 28.05.2015 16:22:31
Too many hyphened lines in a row. This happens unusually often.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:22:48
parameters. For

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:23:00

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:20:38

that enlarge the distance between the similarity of the topic and the similarity of all other objects in the context.

Search is also applied when an agent perceiving an utterance tries to interpret it. The semantic structure an agent parsed from an utterance is often incomplete and semantic entities, cognitive operations and links can be missing in the network. Interpretation is a flexible, active process by which agents use search to add missing items to the network. Networks are immediately evaluated to see if they find a referent for the parsed utterance. The search process is constrained by the partial meaning parsed from the utterance and the kinds of semantic structures that are appropriate in the current context. The same scoring mechanisms as for conceptualization ensures that only structure that is discriminating for a particular object (implicitly assuming that the speaker constructs structure based on these principles) will be considered and the best of all possible results is chosen as the interpretation of an utterance.

3.4 Chunking

Search spaces quickly become intractable because the number of possibilities for composing semantic structures increases exponentially with the number of cognitive operations. A look at language is helpful here. Grammar can be analyzed as a sophisticated tool that highly structures human language in order to manage not only the search space of possible syntactic structure (Steels & Wellens 2006) but perhaps more importantly the vast space of possible conceptual structures. Parts of meaning that are covered by a particular part of language can be stored as a *chunk* and then used as an basic atomic unit in composition. Ready-made semantic structure dramatically reduces the search space. If a structure like the final structure in Figure 3.1 is constructed from scratch using simple operations, the search tree would have a search depth of three (essentially one step in depth per operation). However, every time an operation is added to a program, it can be linked to the current structure in multiple ways, which leads to an explosion of nodes on every layer of depth. Hence, the system soon has to deal with a wide search tree, where every node will be executed and tested against the context. Consequently, using chunking dramatically increases the performance of the system, even in simple examples.

Chunks have an important role in the study of language because they package strategies for conceptualizing reality. Chunks allow to research how cognitive operations form strategies for conceptualizing reality, how these strategies can be adapted by agents and how strategies become conventionalized. For studying

Seite: 40

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:27:33
" A ready-made semantic ... " ?

3.6 Discussion

IRL is a powerful system that for the first time allows to study complex semantic phenomena that go beyond purely lexical studies. IRL is a general system for representing the procedural semantics of utterances. It establishes a link between perception and language by providing a mechanism for representing the meaning of utterances, finding and interpreting the meaning of utterances. Moreover, IRL is designed to allow language processing to be a flexible, adaptive process which can be extended by new cognitive operations, new chunks, and new categories at any moment. Moreover, IRL provides mechanisms for tracking the success of semantic material such as chunks and categories.

The oldest and in some sense most similar system to what I have presented here is Winograd's SHRDLU (Haddock 1989; Winograd 1971), However, SHRDLU misses the key aspects of grounding, active interpretation and conceptualization as a search process. Other work such as Bailey et al. (1997) and Siskind (2001) focus mostly on lexical meaning. Some approaches have taken more general approaches e.g. to event structure (Narayanan 1999) but stay mostly tied with that particular domain. One of the few approaches talking about objects and events in the same framework is Roy (2005), which is comparable to ours, but so far has been a theoretical proposal only.

Seite: 43

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:32:39

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:32:37
It is often a good idea to use small caps for long abbreviations

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:33:51
The LSP guidelines suggest not using the academic we, unless you want to refer to "Steels & Spranger" or s.th. similar

4 Construction Grammar with FCG

Construction Grammar posits that linguistic knowledge is organized in the form of *constructions* (Goldberg 1995; Croft 2001) which are mappings of semantics and pragmatics to syntax, i.e. words and grammar, but also phonology, prosody or intonation. Typically, construction grammarians take a functional view on language and analyze every piece of language as a tool for communication and in terms of the syntactic and semantic function it performs. The theoretical framework of Construction Grammar is important for this book, because it integrates semantics with syntax and opens up ways for understanding the acquisition and evolution of language as a tool for solving communicative problems in which all elements of processing from semantics to syntax can be used as a tool for solving these problems.

Every part of an utterance has meaning and a semantic function. The meaning of a lexical item is the reference to the category, prototype or concept that it refers to. Its function is how it is used both in the semantic structure underlying the phrase and in the syntactic structure of the phrase. The following two examples include the word "rot" (red) but they use the word in completely different syntactic and semantic structures.

1. "der rote Block" (the red block)

2. "Rot ist eine Farbe" (red is a color)

In Example 1 "rot" (red) is used to modify the set of objects denoted by the word "Block" (block) whereas in 2 the statement is about the color itself. We can precisely capture these differences in semantic function using cognitive operations and IRL (an structure for Example 1 can be found in Section 3). The semantic function is coupled to a particular expression in syntax. In Example 1 the color is expressed as an adjective which signals its use as a modifier. In Example 2 the color is expressed as a noun and signals that the subsequent verb phrase is a fact about the color itself. In production, the speaker can therefore choose to express the category as an adjective if the category is linked to the corresponding cognitive operation (e.g. `apply-color`). In parsing, when he observes a color adjective this allows him to infer that he is supposed to modify a set of objects using that

Seite: 45

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:44:12

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:44:19

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:44:34
These examples should be typeset in \ea \ex \z \z

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:44:40

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:45:17
This would become "In (1)" if we adopt the \ea \z code

operation. Which set of objects the color adjective modifies is determined by the larger syntactic and semantic context. For instance, in Example 1 the adjective is part of an adjective noun phrase that indicates which set is modified by the color category namely the set of blocks. From the viewpoint of the adjective noun phrase the adjective has the semantic function of providing a modifier and in particular of modifying the set of objects denoted by the noun. Of course, other adjectives, such as spatial adjectives can have the same function within an adjective noun phrase. The modified set is then input to another operation namely the operation `apply-selector` which is marked by the determiner. So what we can see already in these simple examples are mappings from semantics to syntax and back, where every aspect of syntax, i.e. words and grammatical relations have a specific effect on the semantic interpretation of the phrase. Vice versa, the speaker can use all the potential of syntax to communicate precise semantic distinctions that he wants to convey. The key item for analysis is the function of items both in syntax and semantics. These dependencies between syntax and semantics can be easily operationalized using Fluid Construction Grammar (FCG) (De Beule & Steels 2005; Steels et al. 2005).

Throughout this book language processing is implemented in a computational implementation of Construction Grammar called Fluid Construction Grammar (FCG) . FCG is 1) a formalism that provides a notation for specifying constructions, 2) a an engine that processes linguistic structure by applying constructions, in order, to produce utterances or parse meanings, 3) a set of design principles for organizing the grammar and linking grammar to representations of semantics, in particular, to semantic structure formalized using IRL.

4.1 Linguistic Processing

Linguistic processing encompasses both production and parsing of utterances. In production, FCG starts from a conceptualized meaning and tries to translate as much as possible of the semantic structure conceptualized by IRL into syntactic structure, i.e. words and grammatical relations using constructions in the linguistic inventory. In parsing this process is reversed and the construction inventory is used to recover as much semantic structure from an utterance as possible. Processing is organized around the *transient structure* which acts as a blackboard representing the current state of processing. Constructions work like rules. If a construction is applicable, i.e. if conditions for its application are met, the construction can change the transient structure. Over time the transient structure accumulates information provided by the different constructions that

Seite: 46

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:46:42
extra space

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:47:04
usually in this book, (1) is used for such purposes, not 1)

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:47:26
maybe better: "meanings, and (3) ... "

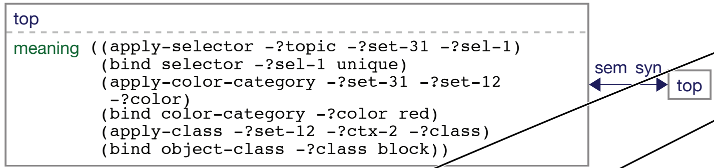

Figure 4.1: Initial transient structure which contains only the meaning to be expressed in the top-unit of the semantic pole (left). There is no hierarchy yet and the syntactic pole (right) is empty.

have applied until some end state is reached, for instance, no construction can apply anymore.

Transient structure The transient structure has two poles: a semantic and a syntactic pole. Information regarding meaning is accumulated on the semantic side, information about words and grammatical relations are gathered on the syntactic side. Information is organized into units identified by a unit-name. Units consist of attribute-value pairs. In order to represent constituent structure, units can form hierarchies in which some units are hierarchically linked to other units effectively building tree like structures. In the beginning of processing the transient structure is filled with information either from the conceptualization processes, e.g. in production, or from the utterance observed, e.g. in parsing. Subsequently, constructions change the transient structure by adding new units, introducing hierarchy, changing the value of attributes or by introducing new attributes. Figures 4.1 and 4.2 show the transition from an initial transient structure which only contains a single unit, called *top-unit* on each side to a final transient structure with hierarchical organization of units and many more features. The initial structure only contains a meaning on the semantic side. The final structure contains, among other things, strings and syntactic word order constraints which can be used to build an utterance, a process called *rendering*. Figures 4.1 and 4.2 show graphical representations of the list representation (s-expression) used in processing. The following restates the initial transient structure as s-expression.

```
((top
  (meaning ((apply-selector -?topic -?set-31 -?sel-1)
            (bind selector -?sel-1 unique)
            (apply-color-category -?set-31 -?set-12 -?color)
            (bind color-category -?color red)
```


Seite: 47

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:48:42
..."in the "top" unit..." ?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:49:06
Odd line spacing in this par. Is this intentional?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:49:43
unit name. Or do you want to refer to a part of code? Then, unit-name in console style font would be better, I think

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:49:49
s.a.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:50:15
s.a.

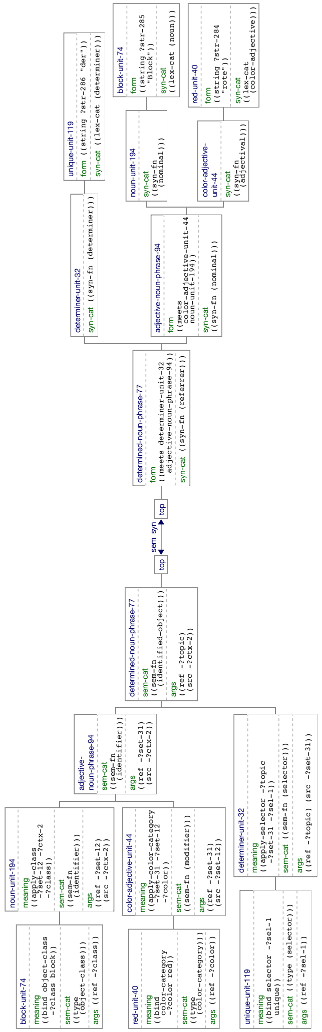

Figure 4.2: Final transient structure after many constructions of a simplified German grammar have applied. The structure consists of units which are hierarchically organized starting from the top-unit. The meaning to be expressed is distributed over various units on the semantic side. Units feature semantic and syntactic categorization (`sem-cat` and `syn-cat`) which was build up in processing to organize constituent structure and allow for high-level constructions to abstract from individual items. On the syntactic side units have `form` features consisting of strings providing words and, so called *meets constraints* which introduce word order.

Seite: 48

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:51:00

```
            (apply-class -?set-12 -?ctx-2 -?class)
            (bind object-class -?class block)))))
<-->
((top))
```

The top shows the semantic pole. The bottom, after ~~the~~ `<->`, shows the syntactic pole. Both poles have one unit (the top-unit). On the semantic side the top-unit has one attribute, the meaning attribute which has as its value an IRL-network in list form. The following shows the final structure in the same representational format.

```
(...
 (color-adjective-unit-44
  (meaning ((apply-color-category -?set-31 -?set-12 -?color)))
  (sem-subunits (red-unit-40))
  (sem-cat ((sem-fn (modifier))))
  (args ((ref -?set-31) (src -?set-12))))
 (red-unit-40
  (meaning ((bind color-category -?color red)))
  (sem-cat ((type (color-category))))
  (args ((ref -?color))))
  ...
 (top (sem-subunits (determined-noun-phrase-77))))
<-->
(...
 (color-adjective-unit-44
  (syn-subunits (red-unit-40))
  (syn-cat ((syn-fn (adjectival)))))
 (red-unit-40
  (form ((string ?str-284 "rote")))
  (syn-cat ((lex-cat (color-adjective)))))
  ...
 (top (syn-subunits (determined-noun-phrase-77))))
```

Only parts of the complete final structure are shown, in particular, three units on each pole are shown the `red-unit-40`, `color-adjective-unit-44` and `top`. In contrast to the initial transient structure, meaning is distributed across different units. Notice that which unit is subunit of another is coded by a special attribute called `syn-subunits` on the syntactic pole and `sem-subunits` on the semantic pole. Compare this with Figure 4.5 which shows the hierarchy in the final structure. For example `red-unit-40` is a subunit of `color-adjective-unit-44`.

Constructions Constructions are organized in the same way as transient structures. They consist of two poles and the data in each pole is organized in terms of units, attributes and values. FCG supports bi-directional constructions which means that the same construction is used in production and parsing. The differ-

Seite: 49

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 16:51:44

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:51:37
Odd line spacing continues here

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:52:07

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:52:11

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:52:12

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:52:45
Maybe this would be better set in a subsection?

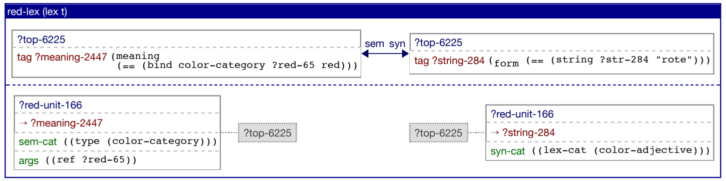

Figure 4.4: Schematic representation of a construction. The two poles of the construction are shown. The top shows the tagged and matching parts of the construction. The bottom shows the hierarchy building part of the construction.

ence between production and parsing is how the syntactic and semantic pole of a construction is used in each case. In production the semantic pole is used to check the applicability of the construction. In parsing the syntactic pole is used. Applicability of a construction is checked using a mechanism called *matching*. Matching is based on the well studied concept of *unification* which is a computational process for equating two terms in this case the semantic or syntactic pole of the construction with the corresponding pole of the transient structure. If matching succeeds, the construction can change both poles of the transient structure, a process called *merge* because it fuses information. The precise inner workings of these two fundamental operations are described in Steels & De Beule (2006). The most important fact is that matching in FCG mainly relies on variables, which in FCG (and in IRL) start with `?`. In computational terms constructions specify 1) under which conditions they apply and 2) if they apply how the structure should be changed.

Figure 4.4 shows an example of a lexical construction that maps the color category `red` onto the string “rote” (red) (Figure 4.5 shows what happens when this construction is applied to the initial transient structure). The following shows the low-level list representation of the construction schematically depicted in Figure 4.4

(4.1)
```
((?top-6143
  (tag ?meaning-2381
   (meaning (== (bind color-category ?red-57 red)))))
 ((j ?red-unit-158 ?top-6143)
```


Seite: 51

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:54:37

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:54:45
s.a. on (2) vs 2)

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:55:15
Should this be a marker for s.th like "code snippet 4.1"

```
   ?meaning-2381
   (sem-cat ((type (color-category))))
   (args ((ref ?red-57)))))
 <-->
 ((?top-6143
   (tag ?string-251 (form (== (string ?str-251 ”rote”)))))
  ((j ?red-unit-158 ?top-6143)
   ?string-251
   (syn-cat ((lex-cat (color-adjective)))))))
```

The top displays the semantic pole followed by the syntactic pole after the `<-->`. In production the construction requires the meaning `(bind color-category ?red red)` to be present. if this is the case the construction merges the information on the syntactic side, in particular, the stem into the transient structure. Additionally, this construction builds hierarchy. It introduces a new unit which is a subunit of the top-unit and which is used to collect information for this particular lexical item. Already this simple construction uses the four basic ways in which constructions interact with the transient structure.

Variables and matching Constructions inevitably contain many variables. Already the unit names in the transient structure are changing every time a new utterance is parsed or a new meaning is produced. But also, just to give another example, variables in the meaning linking cognitive operations are different every time IRL conceptualizes. Using a variable in one part of the construction and repeating it in another can lead to changes in the transient structure triggered by matching and merging (Steels & De Beule 2006). Example 4.1, for instance, uses matching and merging by re-using the variable in the meaning `?red-57` in the `args` attribute. Whatever this variable binds to in processing the re-occurring variable will make sure that the data is available in both places.

Hierarchy Hierarchy is built using a special operator called the *J-operator* which changes the transient structure to include a new unit (De Beule & Steels 2005). The new unit can have units that are already present in the transient structure as children. A construction can therefore easily change the hierarchical structure of the complete pole. The J-operator syntax is the following.

```
((J ?new-unit ?parent (?child-1 .. ?child-n))
 (new-attribute new-attribute-value))
```

In the example construction the J-operator is used on the semantic and on the syntactic side. It introduces new units on both sides and adds information to this unit (in Figure 4.4 the parts pertaining to the J-operator are

Seite: 52

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:55:38
If

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:55:50

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:56:12
Odd line spacing. Is this intentional?

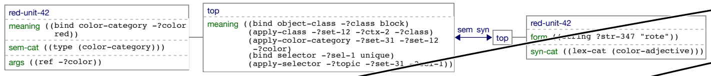

Figure 4.5: Transient structure after the lexical construction applied. The construction has introduced two new units using the J-operator. One on the semantic side and one on the syntactic side. Both units have the same name `red-unit-42`. The construction introduced the string "rote" on the syntactic side and the bind statements used for triggering the construction has been moved moved using the tag-operator from the top-unit to the new semantic subunit. The construction also added new semantic and syntactic categories (`sem-cat` and `syn-cat`) that can be used by subsequent constructions.

shown below the dotted line). Notice that the name of the new units is equal.

Movement Constructions can *tag* attributes and their values, in order, to move them around. In this example construction the tag-operator moves the bind statement pertaining to the color category from the top-unit to the newly created unit. The tag operator takes the following form

```
(?unit (tag ?tag-variable (attribute attribute-value))
```

The operator binds whatever follows the variable `?tag-variable` to the variable. If the variable is used in a J-unit, i.e. a unit with a J-operator, in another part of the construction this denotes the place where `(attribute attribute-value)` will be moved. The example construction has tag operators on the semantic side for moving the bind statement to the new semantic unit. Similarly, on the syntactic side the operator is used to move the string "rote" to the new syntactic unit.

Search Constructions are organized in a pool of constructions. In processing constructions, in principle, compete for access to the transient structure. Typically, more than one construction can apply to the transient structure and the question is how to organize the process if there are multiple constructions that

Seite: 53

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:57:26
I think there is an extra space here between "Figure 4.5." and the beginngin of the desc.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:57:39
construction WAS / IS applied ?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:58:15
Shouldn't this be set in console style text, too?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:58:30
Odd line spacing. Is this intentional?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:59:48
It is ambiguous WHO competes for access.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:59:16
While this par is not indented, the previous two (movement, hierrachy) were.

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 16:59:35

Figure 4.6: FCG search tree which produces “der rote Block” given the IRL-network shown in Figure 3.1.

want to change the transient structure. In the absence of apriori rules to prefer one construction over another, each construction that can apply to the transient structure is tried in a different branch of a heuristics guided search process. In other words, instead of having competing constructions change the same transient structure, the structure is copied and each potentially applying construction is applied to a copy without necessarily influencing the other. Naturally, this leads to different branches in processing in which each branch computes a particular parsing or production result. Search is represented using a search tree in which each node contains a transient structure. The initial node contains the initial transient structure. Leaf nodes contain final structures. The search process itself can be manipulated. For instance, it is possible to remove and refrain from processing duplicate nodes which contain the same transient structure and the order of following a particular branch can be influenced by how successful one predicts the branch to be. Figure 4.6 shows an example search tree for production of the utterance “der rote Block”.

Design Layer In order to design grammars it has proven beneficial to abstract ~~away~~ from the low level processing layer of FCG and add a representational layer that connects high level linguistic analysis with the processing engine of FCG. The idea is to allow re-occurring problems in grammar design to be solved using *templates* without having to resort and copy the code needed for describing a construction in the basic list notation. Templates are a general mechanism for expressing *design patterns*, i.e. solutions that can be re-used to deal with the same problem occurring in different situations. For instance, all grammars implement phrasal constructions. One of the main semantic functions of phrasal constructions is to introduce variable equalities for linking constituents. A template encapsulates the solution to the problem of linking constituents in a way that the solution can be re-used in other phrasal constructions of the same grammar, but ideally of course also for phrasal constructions in other grammars. Templates are defined similar to functions. They have a name and a set of arguments which

Seite: 54

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:00:10
should be set in italics

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:00:21
a priori

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:01:17
should be set in italics

Verfasser: fkopecky Thema: Durchstreichen Datum: 28.05.2015 17:01:27

From a computational perspective, FCG provides an easily manipulatable data representation which makes inventing new constructions, changing and adapting semantic and syntactic categories, introducing hierarchy or movement of data relatively easy. In line with construction grammar the *unified* nature of representation is important. There is absolutely no difference in terms of representation and processing between lexical, functional or phrasal constructions. Hence, FCG supports research into how constructions can change from lexical to grammatical constructions, which is of interest for the study of the influence of grammaticalization processes on language evolution(Traugott & Heine 1991).

Another important argument for the use of FCG is its robust behavior in parsing and production. The search process for construction application and the bi-directional nature of constructions allow agents to produce as much of the meaning as they can when they are speaker. In parsing, the same process allows agents to recover as much of the semantics of a phrase as they possibly can. This is a prerequisite for any kind of grounded language learning let alone language evolution. Agents have to get as much information as possible from the different systems such as perception and conceptualization, but also language processing. If agents would have to deal with a grammar engine that essentially gives up on processing as soon as an agent encounters a phrase that he thinks is unconventional, learning the new unconventional phrase can never occur or is significantly hindered. Whereas if the grammar engine provides as much information as possible agents have a much better shot at guessing underlying meaning and making sense of what was conveyed to them. Subsequently, they can better represent the new parts of an utterance versus parts they already know. Modeling this whole process as a search process is an immense advantage of FCG. Agents can track what changes when they apply other constructions and explore different possible parse and produce results, in order, to identify problems in language processing.

The last point with respect to the advantage of keeping information from the search process that governs linguistic processing is important, in particular, for the main problem studied in this book: conventionalization. In order, for agents to realize that constructions are competing for the same string, the same grammatical structure or the same meaning, it is vital to fully explore the search process. If there are multiple ways of producing an utterance for a meaning, for instance, because there are multiple words to express the same category, then the search can recover all of them. Together with a mechanism for tracking success of constructions, the search can choose the best one of them. After the interaction the agent can then update the constructions used and those that he could

Seite: 56

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:04:40
are the speaker ?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:06:01
order for ?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:06:32

have used, for instance, by rewarding successfully used constructions and punishing unsuccessful or unused constructions. Constructions are equipped with a score that allows agents to update their inventories by scoring constructions according to their success in communication. If scores get too low agents can remove the affected constructions.

4.3 Discussion

There is no question that this is a short, in many ways too short, introduction to FCG. FCG has been under continuous development since 1998 and it has developed into a mature system which allows to to research complex language phenomena such as Russian aspect (Gerasymova & Spranger 2010, 2012). The complexity of natural language has no doubt left its mark on the system and many design choices in the system are not immediately obvious, unless one takes the scope of the research program into account. Recently several book projects (Steels 2011b, 2012) attempted to communicate the full scope of FCG research performed in the last 5-10 years. The interested reader is referred to these efforts to get a broader introduction.

Seite: 57

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:07:06
"has without doubt" ?

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:07:37
please make sure this is an en-dash

5 German Locative Phrases – An Introduction

To appreciate the complexity of spatial language one just needs to consider a particular human language such as German spatial language. The following chapters detail an elaborate reconstruction effort which targets locative German phrases (parts of this reconstruction effort have been published in Spranger & Loetzsch 2011). I specifically focus on the processing of German locative phrases in a whole systems approach encompassing the perception, conceptualization, as well as production and parsing of spatial phrases. Before we jump to the implementation and the specific challenges in modeling such a complex phenomenon, this introduction overviews German spatial relations and highlights the syntax and semantics of German locative phrases, as well as the close interaction of syntax and semantics. The claim is that important aspects of the syntactic structure of an utterance, i.e., the lexical items and the grammatical relations between them, work together to convey semantic structure, i.e. meaning. Vice versa, the varied syntactic devices in German spatial language allow to express subtle differences in the conceptualization of spatial scenes. The German spatial language system serves as a beautiful example of how syntax connects to the extraordinarily rich world of spatial semantics. The following paragraphs give a quick introduction to German locative expressions.

The literature distinguishes several classes of spatial relations available in German. Among them are projective, proximal and absolute relations (see Figure 5).

Projective Relations sometimes also called *dimensional* terms (Eschenbach 2004; Herskovits 1986) in German comprise the class of six items referring to spatial dimensions "vor" (front), hinter (behind), "über" (above), "unter" (below), "rechts" (right), "links" (left) (Tenbrink 2007, 2005b; Wunderlich & Herweg 1991). Traditionally, and for reasons of distinct syntax and semantics the class of projective relations is further divided into *frontal* ("vor" and "hinter"), *lateral* ("links" and "rechts"), *horizontal* (comprising lateral and frontal relations), and *vertical* relations ("über" and "unter").

Proximal Relations are part of the larger class of topological relations that struc-

Seite: 61

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:09:24

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:09:35

ture space with respect to proximity, contact and inclusion (Grabowski & Weiss 1996). For this book proximal relationships such as "nah" (near) and "fern" (far) are important.

Absolute Relations refer to cardinal directions, for instance "nördlich" (north), "westlich" (west), "östlich" (east) and "südlich" (south).

Spatial relations take different syntactic forms in German. All of the projective terms, for instance, can be expressed in different lexical classes, most prominently as adjectives, adverbs and prepositions. For example, the projective term "vor" can appear as adjective as in Example 5.1, as adverb as in Example 5.2 and as preposition as in Example 5.3.

(5.1) *der* *vordere* *Block*
the.NOM front.ADJ.NOM block.NOM
'The front block',

(5.2) *der* *Block* *vorne*
the block.NOM in the front.ADV
'The front block'

(5.3) *der* *Block* *vor* *der* *Kiste*
the.NOM block.NOM front.PREP the.DAT box.DAT
'The block in front of the box'.

The different lexical classes carry with them different syntactic functions, e.g., adjectives can function as modifiers in determined adjective noun phrases, prepositions are followed by noun phrases and in German govern case. But, each lexical class is also connected to a different semantic interpretation. In particular, there is a tight connection between the lexical class and specific spatial construal operations that govern how precisely the spatial relation is to be applied. The meaning of a projective category when used as an adjective is to filter objects (Tenbrink 2007), whereas when used as preposition the meaning is to construct a region (Klabunde 1999). The following phrase, for instance, uses the projective category front to construct a region to which one is asked to put a chair. Unlike in the adjective case, the region is not used to modify or filter, rather the region is necessarily empty, in order, for the chair to be put there.

(5.4) *"Stelle* *den* *Stuhl* *vor* *den* *Schrank!"*
"Put the chair front.PREP the cupboard"
"put the chair in front of the cupboard"

Seite: 62

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:09:52

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:09:50

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:10:26
General remark: We need to typeset the category abb. in small caps

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:10:31

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:10:35

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:10:36

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:11:11

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:11:14

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:11:17

Another important difference between spatial adjectives on the one hand and spatial adverbs and prepositions on the other, is the potential for *reference objects* or *landmarks*. Example 5.3 relates a *located object*, also called figure (Talmy 2000) and trajector, (Vandeloise 1991), to a *reference object*, also called landmark (Vandeloise 1991), relatum (Tenbrink 2007) or ground (Talmy 2000). If a spatial term is used prepositionally, as in Example 5.3, both the spatial relation, in this case "vor", and the landmark "Kiste" (box), are expressed in the prepositional phrase itself.

A third possible lexical class in which spatial relations partake are adverbs. Spatial adverbs can be accompanied by a prepositional phrase, as in the following Examples:

(5.5)	*der*	*Block*	*vorne*	*in*	*der*	*Kiste*
	the.NOM	block.NOM	front.ADV	of.PREP	the.DAT	box.DAT

'The block in the front of the box'

(5.6)	*der*	*Block*	*links*	*von*	*der*	*Kiste*
	the.NOM	block.NOM	left.ADV	of.PREP	the.DAT	box.DAT

'The block to the left of the box'

The prepositional phrases "in der Kiste" and "von der Kiste" both introduce a landmark which relates to the figure, in this case "der Block", via the spatial term. The two different prepositions "in" and "von" denote whether the relation referred to by the spatial term is *internal* to the landmark or *external*. In the case of "von", the spatial region denoted by the projective adverb, e.g., "links" (left), is external to the landmark, whereas in the case of "in" the region lies within the landmark. The projective adverbs "links" (left) and "rechts" (right), can be followed both by "von" and "in" prepositional phrases, hence, they can have an internal and external reading. The projective adverbs, "vorne" (front) and "hinten" (back) can only be extended by "in" prepositional phrases. The vertical projective adverbs "oben" (above) and "unten" (below) elicit internal readings. Again, differences in semantic processing are syntactically marked, in the case of adverbs by prepositional phrases that complement the adverb.

The last important component of spatial language considered in this book is perspective marking. The following two example feature perspective markers, in the first example an adverb is perspective marked, in the second a prepositional phrase is perspective marked.

(5.7)	*der*	*Block*	*vorne*	*von*	*dir*	*aus*
	the.NOM	block.NOM	front.ADV	from.PREP	the.DAT	box.DAT

Seite: 64

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:11:55
superfluous ,

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:11:46

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:12:15

‘The block in front from your perspective’

(5.8) *der* *Block* *links* *der* *Kiste* *von*
the.NOM block.NOM left.PREP the.GEN box.GEN from.PREP
dir *aus*
your.DAT perspective
‘The block to the left of the box from your perspective’

Perspective on a scene is important for particular interpretations of spatial phrases because it influences how the spatial scene and, in particular, how the landmark is conceptualized.

These few examples from German locative phrases show that we can analyze the syntax of spatial language fruitfully in terms of what is its spatial semantics. This resonates with theories of syntax such as construction grammar, which put the direct mapping of syntax to semantics at the core of language processing. The tight relationship between syntax and semantics is an important claim in this book which underlies the reconstruction efforts, and also the evolution experiments.

The first question we are focussing on is how to organize language processing. That is, how semantics is encoded in words and grammar - and vice versa. However, this is not the full story. We need to ground language in perception. Speakers need to be able to plan what they are going to say based on their communicative intention. Hearers must have machinery for interpreting the utterance based on the spatial context. This widens the question from how to organize the syntax and semantics interface to how to organize processing in a large array of systems comprising perception, conceptualization, interpretation and, also but not only, processing of syntax. In particular, one has to identify the cognitive operations underlying the semantics of German locative phrases and how these operations are used in conceptualization of spatial scenes, as well as how conceptualization interacts with the perception of spatial scenes. The following sections examine three main questions. The first is, what is the meaning of phrases such as, for example, “der Block links von der Box von dir aus” (the block left of the box from your perspective, see Example 5.8) and how can we formalize the semantic structure of these utterances. In particular what are the cognitive primitives that are necessary for modeling the semantics of spatial phrases. Second, how does semantic structure get translated into words and grammatical relations and back. And thirdly how can agents autonomously conceptualize spatial scenes given communicative goals and how can agents interpret semantic structure recovered in parsing so that success in communication can be achieved.

Seite: 65

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:14:04
please make sure this is an en-dash

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:14:34

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:14:44
"utterances?"

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:14:58
"phrases?"

Verfasser: fkopecky Thema: Hervorheben Datum: 28.05.2015 17:15:08
"back?"