The evolution of grounded spatial language

Michael Spranger

DRAFT
of Wednesday 23rd September, 2015, 15:39

Computational Models of Language
Evolution 5

language
science
press

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.

DRAFT
of Wednesday 23rd September 2015, 15:3

ISBN 978-3-946234-14-2

Michael Spranger

The evolution of grounded spatial language

Computational Models of Language Evolution

Editors: Luc Steels, Remi van Trijp

In this series:

1. Steels, Luc. The Talking Heads Experiment: Origins of words and meanings.

2. Vogt, Paul. How mobile robots can self-organize a vocabulary.

3. Bleys, Joris. Language strategies for the domain of colour.

4. van Trijp, Remi. The evolution of case grammar.

5. Spranger, Michael. The evolution of grounded spatial language.

ISSN: 2364-7809

The evolution of grounded spatial language

Michael Spranger

DRAFT
of Wednesday 23rd September, 2015, 15:39

language
science
press

Michael Spranger. 2015. *The evolution of grounded spatial language* (Computational Models of Language Evolution 5). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/53
© 2015, Michael Spranger
Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):
http://creativecommons.org/licenses/by/4.0/
ISBN: 978-3-946234-14-2 (Digital)
978-3-946234-15-9 (Hardcover)
978-3-944675-46-6 (Softcover)
ISSN: 2364-7809

Cover and concept of design: Ulrike Harbort
Fonts: Linux Libertine, Arimo
Typesetting software: XƎLaTeX

Language Science Press
Habelschwerdter Allee 45
14195 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

Freie Universität Berlin

Language Science Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate. Information regarding prices, travel timetables and other factual information given in this work are correct at the time of first publication but Language Science Press does not guarantee the accuracy of such information thereafter.

Preface

This book contributes to our understanding of the origins of spatial language by carrying out language game experiments with artificial agents instantiated as humanoid robots. It tests the theory of language evolution by linguistic selection, which states that language emerges through a cultural process based on the recruitment of various cognitive capacities in the service of language. Agents generate possible paradigmatic choices in their language systems and explore different language strategies. Which ones survive and dominate depends on linguistic selection criteria, such as expressive adequacy with respect to the ecological challenges and conditions in the environment, minimization of cognitive effort, and communicative success.

To anchor this case study in empirical phenomena, the book reconstructs the syntax and semantics of German spatial language, in particular German locative phrases. Syntactic processing is organized using Fluid Construction Grammar (FCG), a computational formalism for representing linguistic knowledge. For the semantics the book focusses in particular on proximal, projective and absolute spatial categories as well as perspective, perspective reversal and frame of reference. The semantic investigations use the perspective of Embodied Cognitive Semantics. The spatial semantics is grounded in the sensory-motor experiences of the robot and made compositional by using the Incremental Recruitment Language (IRL) developed for this purpose. The complete reconstructed system allows humanoid robots to communicate successfully and efficiently using the German locative system and provides a performance base line. The reconstruction shows that the computational formalisms, i.e. FCG and IRL, are sufficient for tackling complex natural language phenomena. Moreover, the reconstruction efforts reveal the tight interaction of syntax and semantics in German locative phrases.

The second part of the book concentrates on the evolution of spatial language. First the focus is on the formation and acquisition of spatial language by proposing strategies in the form of invention, adoption, and alignment operators. The book shows the adequacy of these strategies in acquisition experiments in which some agents act as learners and others as tutors. It shows next in language formation 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.

Acknowledgements

This file was included in the folder, but not included. Should we include it?

Certainly, this thesis would not have been possible without a wide range of stable usable software systems such as the Babel2 framework and Fluid Construction Grammar (FCG) used in the experiments. I thank all people involved in developing these systems. 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 these systems: Joris Bleys, Joachim De Beule, Martin Loetzsch, Remi van Trijp and Pieter Wellens. It has been outstanding to work with all of you.

Working with real robots is a tedious but fun task which I have much enjoyed because of the collaboration with Martin Loetzsch and Remi van Trijp in Tokyo. Martin Loetzsch also contributed substantially to the vision system used in the experiments. The robots used in the experiments were graciously provided by Masahiro Fujita, Hideki Shimomura and their team at Sony Corporation in Tokyo. I am grateful for their support.

As part of this thesis, the Incremental Recruitment Language (IRL) was developed and enhanced. There are number of people who have worked on IRL. Wouter van den Broeck developed an early version of the system. Joris Bleys, Martin Loetzsch and Simon Pauw provided many ideas and implementations making the system as usable as it is today. I thank all of them for their help and collaboration.

Martin Loetzsch and Simon Pauw had big impact on many issues discussed in this thesis. I am indebted to both 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. Above all, I thank both of them for the fun and excitement that characterize the last three years of my life.

There are a number concrete contributions that helped to make this thesis. I thank Kateryna Gerasymova and Nancy Chang for their helpful comments on earlier versions of the thesis. I thank Joris Bleys for providing the TEX-macro for semantic operations and Remi van Trijp for his translation of the summary into the funny Flemish language. Moreover, I much appreciate Lara Mennes help in dealing with the university administration and Katrien Beuls great efforts in getting the thesis printed when I was not in Brussels.

I am indebted to the members of my PhD committee, in particular, Tony Belpaeme, Thora Tenbrink as external reviewers and Tom Lenaerts, Bernard Manderick, Anne Nowe and Beat Signer as internal reviewers who have provided valuable comments and helped improve this work substantially. I am extremely grateful to Luc Steels for supervising this thesis and mentoring my intellectual development. Most if not all fundamental ideas put forward in this thesis are based on his extraordinarily rich and exciting work.

I would also like to thank my family, especially my parents, grandparents, my brother and his family for their immense and unconditional support. Finally, I could not have become somebody who finishes a thesis without Katya.

The research reported in this thesis has been financially supported by Sony CSL Paris and Tokyo as well as the ECAgents (FP6) and ALEAR (FP7) projects.

Contents

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) der Block rechts der Kiste von dir aus
the.NOM block.NOM right.PREP the.GEN box.GEN from.PREP 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).

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 between 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 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 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 (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 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 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.

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 grammatical 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 consist of different conceptualization strategies that have 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* 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* and *co-evolution* of syntax and semantics are used as theoretical pillars.

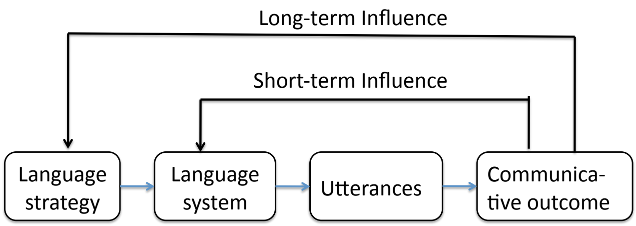

Figure 1.1: The fitness of utterances for communication affects both the language system and the language strategy. The effect of the success of a single utterance on the language strategy is smaller which leads to slower change on the level of the strategy. (Figure adapted from Steels 2011c)

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 a 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 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 they speak and they can adjust their 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 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 basis 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 of the language system under consideration co-evolve over time?

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 to 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 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.

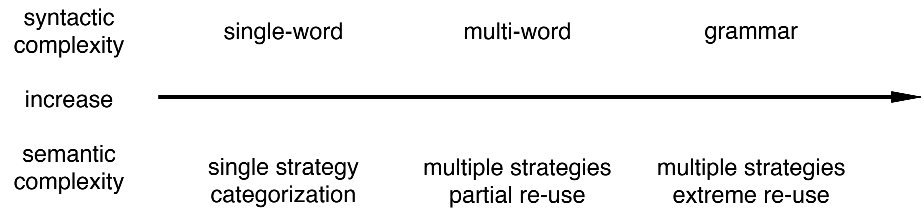

Figure 1.2: Co-evolution of syntactic and semantic complexity.

This book orients itself alongside Steels (2005), who proposes a number of stages of complexity which are relevant for this book: *single-word utterances*, *multi-word utterances*, and *grammatical utterances* (see 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 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 to solve a problem in communication. A speaker, for instance, who is unable to discriminate an object might introduce a new spatial category to be able to identify the object. Subsequently, he might invent a new word to be able to express the new spatial category. Invention operators introduce variation and novelty into the system.

Adoption operators Adoption is the process by which an agent acquires a new word, a new spatial relation or a new piece of grammar. Acquisition is carried out by hearers in interactions when they observe new items that they are unable to process. Adoption is another source of novelty and variation. An agent that picks up a new word might have a different idea of what that word means than the speaker actually intended.

Alignment operators Invention is local to an interaction. When two agents communicate and one of them invents a new word, this word might be acquired by the interlocutor, but the knowledge about this word is still local. Alignment operators orchestrate the self-organization of the system and the global alignment of language.

1.4.2 Co-evolution of syntactic and semantic complexity

In each stage, syntactic complexity co-evolves with semantic complexity (see Figure 1.2). Syntactic complexity rises because the number of words per utterance increases (from the single-word stage to the multi-word stage) and because syntactic categorizations such as word order, morphology, agreement become important (from multi-word to grammar).

The notion of semantic complexity is harder to define. Obviously German spatial language is complex. But why does this seem obvious? What are the properties that make it a complex semantic system? For this book, complex semantics is defined with respect to spatial language as: the language supports a large number of conceptualizations of a spatial scene. There are two factors influencing the complexity of the space of possible conceptualizations of a spatial scene.

Number of relations A first level of semantic complexity is related to the number of spatial categories. For the part of German locative phrases considered in this book, we already have 12 spatial relations. But there are, of course, many more relations not considered in this book such as dynamic relations. For some scholars this is the only definition of semantic complexity (compare Schoenemann 1999).

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 parts besides this introduction and the conclusion. Part I explains the interaction model and the technical systems needed for studying spatial language. Part II deals with objective number one and details the reconstruction efforts for the German locative system. In 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

1.5.1.1 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.

1.5.1.2 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.

1.5.1.3 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 modeling 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.

1.5.2.1 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.

1.5.2.2 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.

1.5.2.3 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 **??** deals with the problem of conceptualization which is the problem of how to construct semantic structure that is helpful in reaching communicative intentions. The chapter gives an overview of different factors influencing the conceptualization of spatial scenes and compares different implementations of spatial conceptualization.

1.5.2.4 Integrating syntactic and semantic processing

The last chapter of this part reports on the integration of syntax, semantics and conceptualization. One of the issues that can be studied in an approach like mine is *semantic ambiguity* which refers to the fact that natural language is often ambiguous with respect to the precise interpretation of a phrase. But humans are very strong in communicating even though language only encodes hints at how to conceptualize reality. The key is that humans integrate the sparse information communicated in utterances with knowledge about the current context of the interaction. Chapter **??** explains how one can operationalize this process of disambiguation through the context using the conglomerate of systems for linguistic and semantic processing as well as perception.

1.5.3 Part III: Spatial language evolution

Finally the book turns to evolution in the third part. The organization of this part orients itself along the stages of complexity introduced earlier. There are two parts on single-word utterance systems, followed by a chapter on multi-word utterance systems. The part closes with a chapter on the evolution of grammatical structure.

1.5.3.1 Acquisition and formation of basic spatial category systems

The first chapter in this part explains how the basic building blocks of spatial language – spatial relationships and corresponding words – become shared in populations of agents. This corresponds to complexity stage one – single words. The goal of the chapter is to define the language strategies necessary for forming single-word spatial language systems.

Single-word spatial language systems are built by a particular strategy of conceptualizing reality which includes a priori commitments to certain reference objects, frames of reference and perspectives on the scene. The chapter shows how a language strategy which is a combination of a particular strategy for conceptualizing reality plus the necessary invention operators for basic spatial categories build the language systems that allow agents to communicate successfully. 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 they use 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.

1.5.3.2 Origins and alignment of spatial conceptualization strategies

Chapter **??** 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.

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.

1.5.3.3 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 **??** 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.

1.5.3.4 Grammar as a tool for disambiguating spatial phrases

The part on language evolution of this book is concluded by 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 ?? 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 ?? 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.

Part I

Spatial language games and technical background

2 Grounded spatial language games

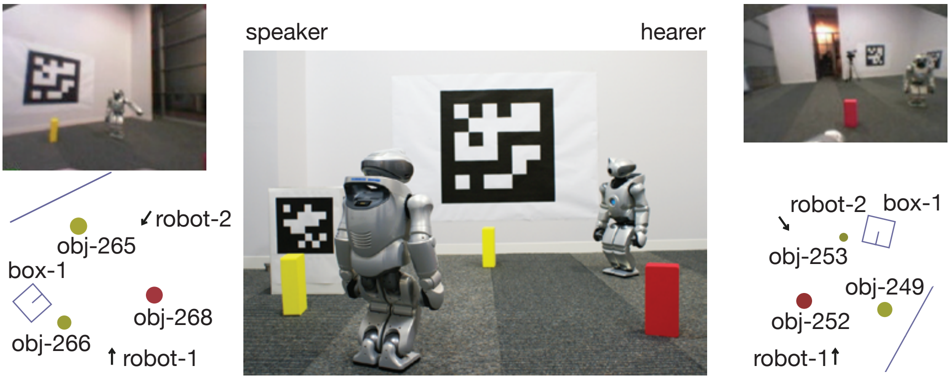

Figure 2.1: Example scene. Two robots autonomously perceive and act in an office environment that contains different types of objects. Both robots autonomously create world models reflecting the state of the environment (see bottom left and right schematics), that include objects with spatial and color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

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 GAMES (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.

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 the language system. The hearer parses the utterance to determine its meaning and interprets it with respect to his current model of the world in order to infer the speaker's communicative goal and, for instance, perform a desired action.

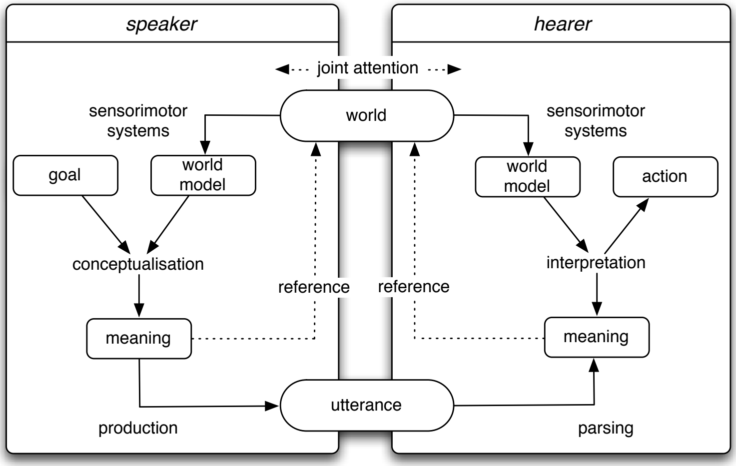

Figure 2.2: The semiotic cycle is a model of situated communicative interactions between two interacting agents.

The system used for conceptualization and interpretation are explained in detail in Chapter 3. The system for producing and parsing utterances is treated in Chapter 4. The following sections focus on the necessary prerequisites for language production, parsing and evolution in terms of social mechanisms and perceptual processing.

2.1 Perception

The environment that robots interact in is equipped with four kinds of objects that were carefully chosen to pick out special features relevant for spatial language: blocks, boxes, wall markers and the robots themselves (see Figure 2.1).

Blocks Are the colored brick like objects. They are typically of the same color. 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 – these are connected regions that differ from the background of the environment or are related to the patterns distributed on the boxes and wall markers. Second, these regions are tracked in subsequent camera images. In order to do so, the vision system needs to establish a correspondence between an internal MODEL and the image regions that refer to the same physical object, a process known in robotics as ANCHORING (Coradeschi & Saffiotti 2003). I use state estimation techniques from robotics, e.g. Kalman ilters (Kalman 1960), for maintaining such persistent models. Third, the vision system fuses information from the proprioceptive sensors of the robot the visual information to encode a set of visual properties about each object. In this particular setup these properties are the position and orientation of objects, an estimated width and height and color information. In the experiments discussed in this book only position and orientation are relevant. Most importantly, the perceptual systems only track objects on the ground. The position and orientation of objects is encoded in a two dimensional *egocentric* coordinate system which has its origin between the two feet of the robot facing to the front of the robot. Spranger (2008) gives more detail on the perceptual systems.

The experiments reported on in later sections require that agents play many language games. In order to speed up the process of a game and in order to do repeatable, manipulatable experiments, data from spatial scenes such as the one in Figure 2.1 are recorded and stored. The output of the perception system of more than 800 spatial scenes with different spatial configurations has been collected and can be accessed by artificial software agents without the agents required to run on physical robots. A spatial language game can be enacted on such stored scenes as if robots were perceiving the scene at the very moment they are playing a particular language game.

Figure 2.3 shows different spatial scenes. Each spatial scene consists of a world model for each of the two robots recorded from the position of each robot. Scenes are grouped into data sets with similar characteristics with respect to perspective on the scene, the number of objects and the availability of boxes and wall markers. For instance, in some data sets the perspective of interlocutors is similar (see Figure 2.3 for examples from one data set), i.e., robots are looking at the scene from the same position and there are few objects. Figure 2.4 shows examples from different data sets.

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 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 take place.

In joint attentional scenes (Tomasello 1995), interlocutors are jointly attending to some object for some reasonable amount of time. Establishing joint attention

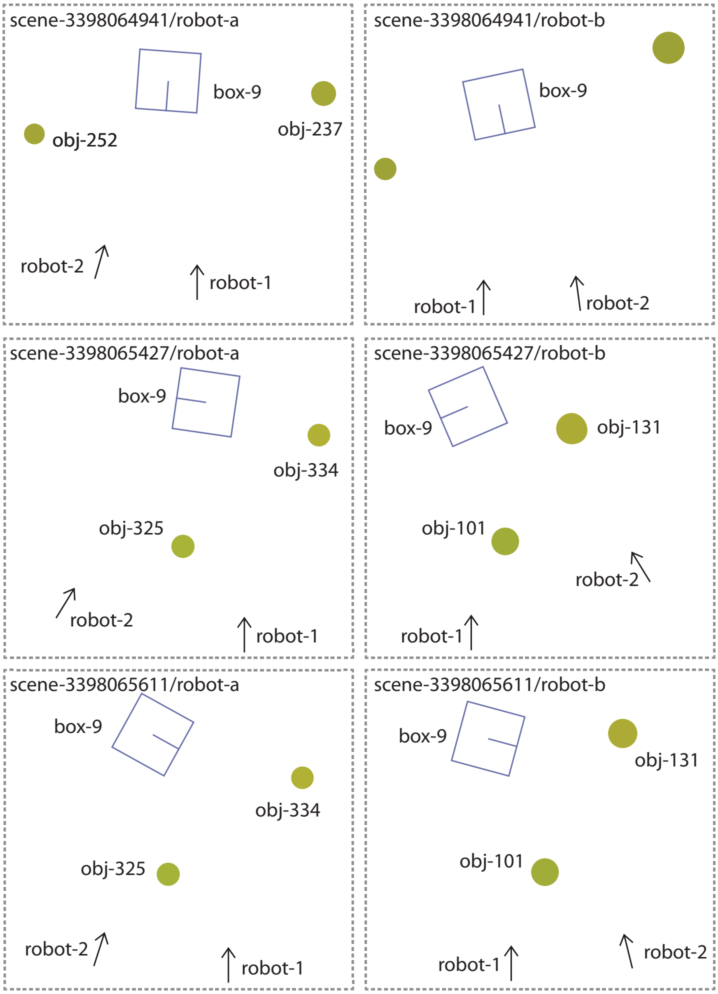

Figure 2.3: Example world models from a spatial data set called *space-game-2*. Left the world model of robot *a* is shown. To the right the world model of robot *b* is shown. All scenes share similar properties. In this data set robots share a similar perspective on the scene. The actual position of the robots varies across different scenes, but is always similar. Similarly, every scene has a box landmark and two yellow blocks in it. However, the actual position of the box and its orientation, as well as the position of the objects change in every scene.

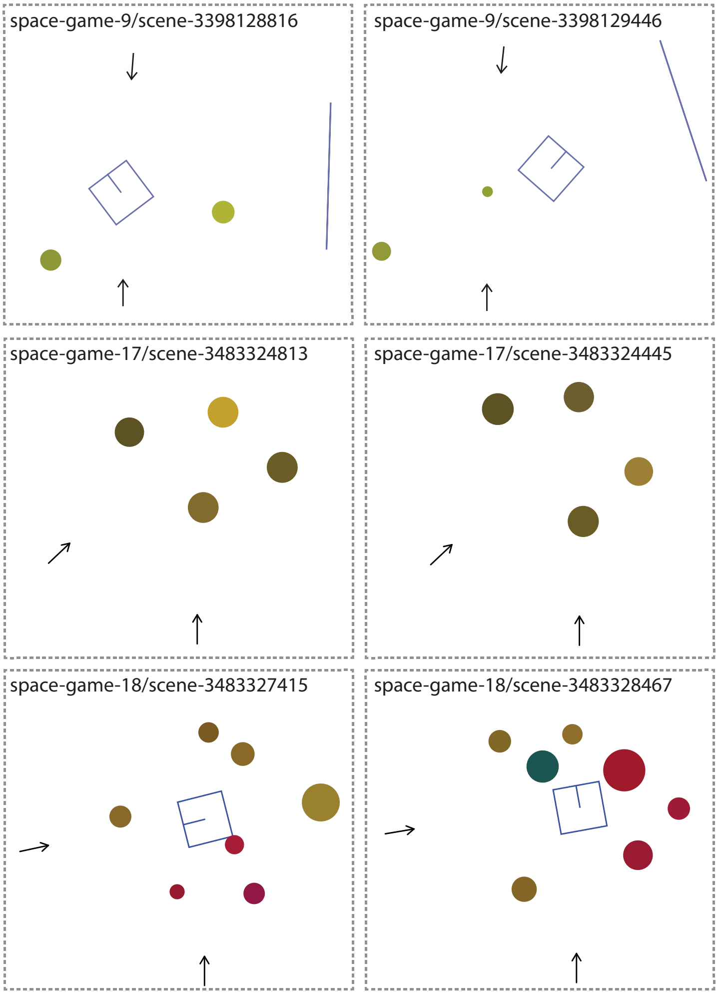

Figure 2.4: Example scenes from different spatial data sets. Each row shows scenes from a particular data set (the world model of robot *a* is always shown). The first row shows a data set which features a global landmark (*space-game-9*). The middle row shows scenes of a data set without global landmark and without box (*space-game-17*). Lastly, a data set which features many objects is shown (*space-game-18*).

in robotic experiments means 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 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.

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 is based on substantial previous work. Key ideas of the IRL system have been laid out by Steels (2000a), with progress reported by Steels & Bleys (2005), Van den Broeck (2008), and recently by Spranger et al. (2010, 2012b).

3.1 Procedural semantics

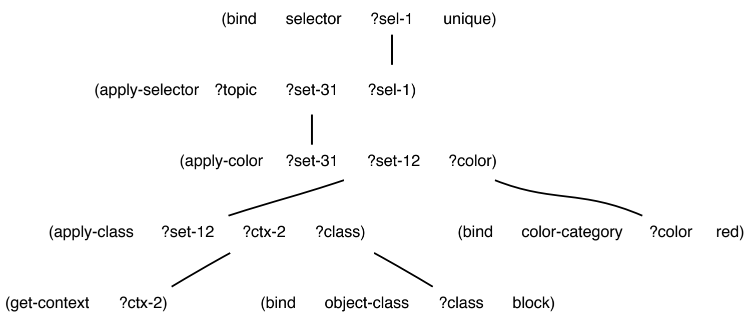

Figure 3.1: Semantic structure underlying the utterance *der rote Block* ('the red block').

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 filter the context for blocks first, followed by the application of the 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 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 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.

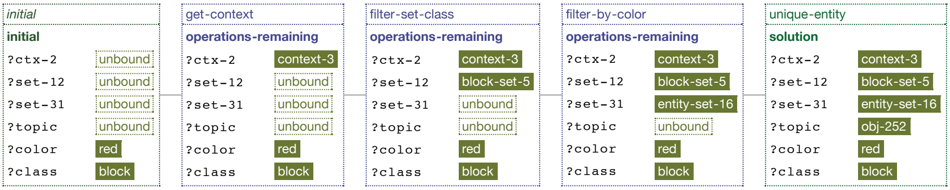

Figure 3.2: Progressive evaluation of the network in Figure 3.1 on the context shown in Figure 2.1. From left to right, each node represents a step in the evaluation process. From top to bottom, the evaluated operation, the node status, and the current list of bindings of each node are shown. A consistent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

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 every object in the context with respect to the object class `block`. This yields the set of objects 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

[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.

bound when the IRL-network is evaluated.

This is only one example how such a network can be evaluated. As Steels (2000a) has argued, 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

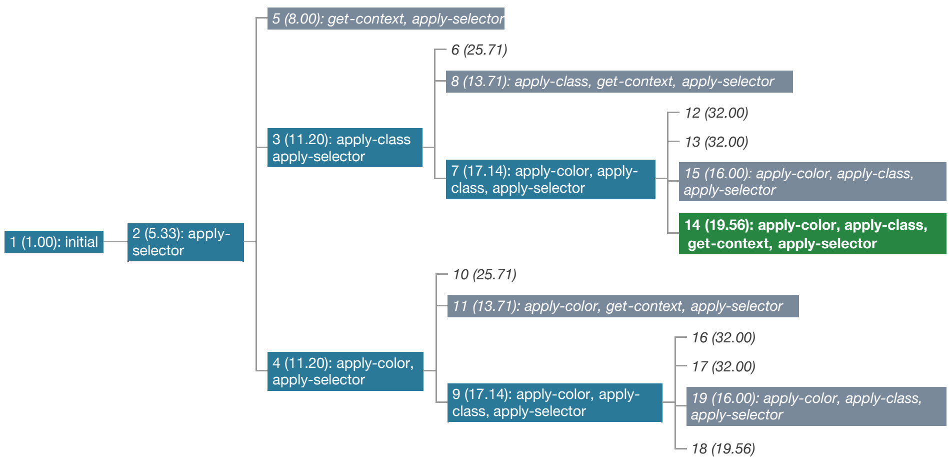

Figure 3.3: The search tree for finding the semantic structure seen in Figure 3.1. From left to right, nodes represent progressively growing programs combined from several chunks, which are each tried out and in some cases lead to solutions (green nodes).

There are two scenarios in which agents autonomously compose semantic structure like the one just described in Figure 3.1. First, speakers have a particular communicative goal and need to construct semantic structure, for instance, for singling out the particular topic they want 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 building semantic structure as a heuristically guided search process, that explores the space of possible IRL-networks driven by the agent's particular communicative goal and the information available to him.

In conceptualization, in other words while "planning what to say" (Steels & Bleys 2005), a speaker searches for an IRL-network that, when executed by the hearer, will reach a particular given communicative goal in a particular context. IRL-networks are constructed by assembling basic building blocks, in particular, cognitive operations packaged into chunks into more and more complex semantic structures. Each assembled structure is immediately tested by evaluating it which assesses its compatibility with the current communicative goal and the perceived context. Figure 3.3 shows an example of such a search process that has produced the program in Figure 3.1 for discriminating the red block in Figure 2.1. The search process for "good" semantic structure is guided by many different heuristics, one being that the structure can be expressed using the language system available to an agent. Others are more focused on the particular character of the communicative goal. If the goal is to discriminate an object in the environment, then it is beneficial to use more discriminative categories, i.e. categories 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 a 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 conventionalization, chunks have scores. This allows to track the success of each strategy with respect to the communicative goals, e.g. discriminative power of a strategy, but also with respect to communicative success. Strategies for conceptualizing reality are tightly linked to language. For instance, one can image that the chunk in Figure 3.4 which is extracted from the network in Figure 3.1 could be the semantics expressed by a determined adjective noun phrase construction. One important claim is that the structure in language – in particular in grammar – is tightly connected to the conceptualization of reality underlying every utterance. So the exemplary facts that in English, noun phrases have determiners, or that in Russian, all verbs are marked for aspects, suggests that these languages require speakers to conceptualize reality in a certain way and mark these conceptualizations in language.

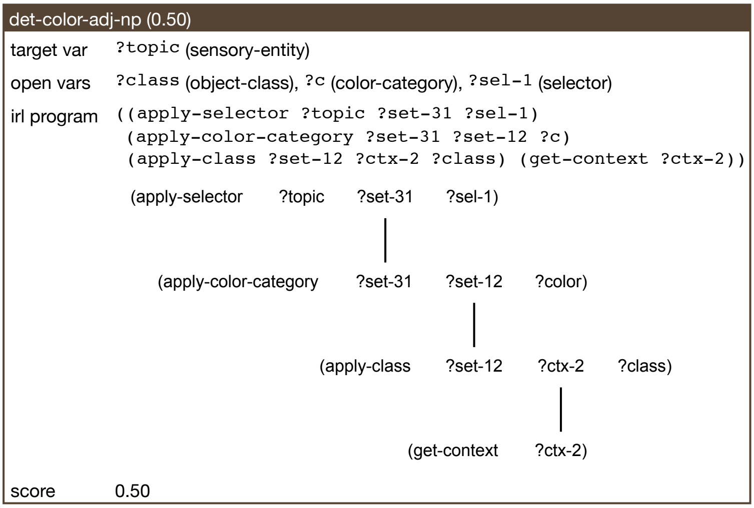

Figure 3.4: Chunk representing the meaning of a determined color-adjective noun phrase. Chunks consist of an IRL-network, plus additional information used for processing: a target variable and open variables. These variables are typed (see brackets for type information) and they are used in conceptualization and interpretation for combining chunks to larger structures. The target variable of a chunk can be linked to the open variable of another chunk. Which selector is used, which object class and which color category are, is mostly determined by evaluating the network which yields bindings for the variables. Since the corresponding variables are open variables the information can also be provided by other chunks or in interpretation by the actual lexical items observed in the utterance. Chunks also have a score which can reflect the degree to which they are conventional ways of conceptualizing reality.

3.5 Grounding

Another important issue is grounding. There are now many proposals of how agents can ground lexicons and categorical systems in sensorimotor interaction with the environment (Billard & Dautenhahn 1998; Vogt 2002; Steels 2008) and IRL is designed to allow such insights to be applied straightforwardly. For instance, the implementation of the operation for `apply-color` is in part based on findings about how basic color categories can be grounded in the sensor data streams of digital cameras (Steels & Belpaeme 2005; Bleys et al. 2009). Similarly, other grounding mechanisms such as for events (Siskind 2001; Steels & Baillie 2003) are easily instantiated in IRL operations.

One of the main claims in this book is that agents co-evolve syntax and semantics. Chunks are one way in which agents can shape strategies for conceptualizing reality. Another is related to the semantic entities themselves and the fact that the number of prototypes and categories and their particular representation is not fixed. For instance, there is now abundant research in the formation of basic color categories (Steels & Belpaeme 2005; Belpaeme & Bleys 2007) and how agents can invent, adopt and shape their inventory of color categories based on the environment they are facing. These insights into adaptive categories, but also names and individuals can be incorporated into IRL, which provides mechanisms for the creation and adaptation of categories in semantic structure.

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 those by 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's (2005), which is comparable to ours, but so far has been a theoretical proposal only.

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 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 FCG (De Beule & Steels 2005; Steels et al. 2005).

Throughout this book language processing is implemented in FCG, a computational implementation of Construction Grammar. 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, and (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 have applied until some end state is reached, for instance, no construction can apply anymore.

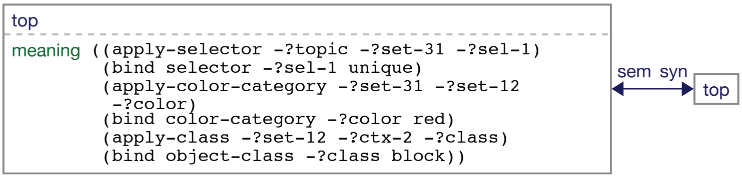

Figure 4.1: Initial transient structure which contains only the meaning to be expressed in the top-unit of the semantic pole (left). There is no hierarchy yet and the syntactic pole (right) is empty.

4.1.1 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
```

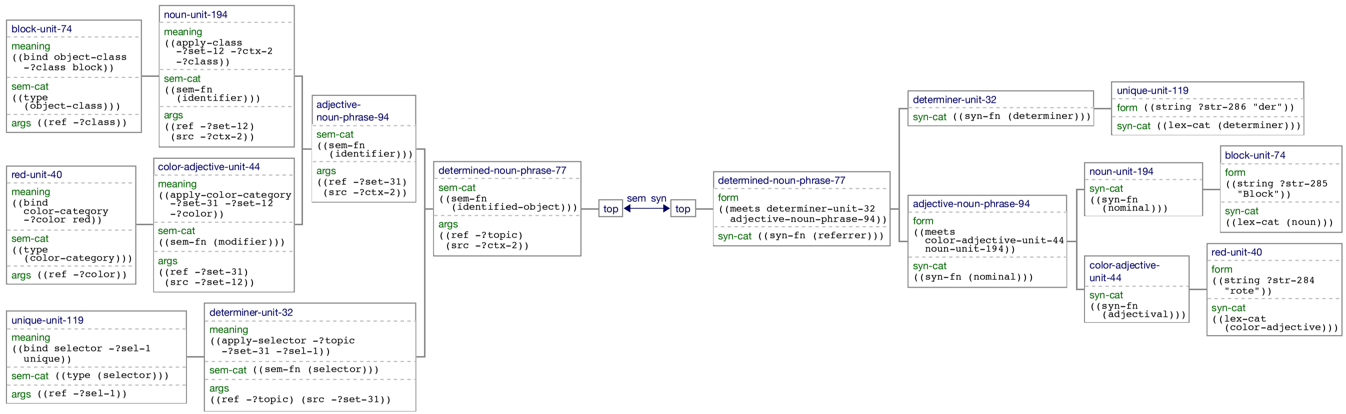

Figure 4.2: Final transient structure after many constructions of a simplified German grammar have been applied. The structure consists of units which are hierarchically organized starting from the top-unit. The meaning to be expressed is distributed over various units on the semantic side. Units feature 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.

```
  (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)
            (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 an IRL-network in list form as its value. 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`.

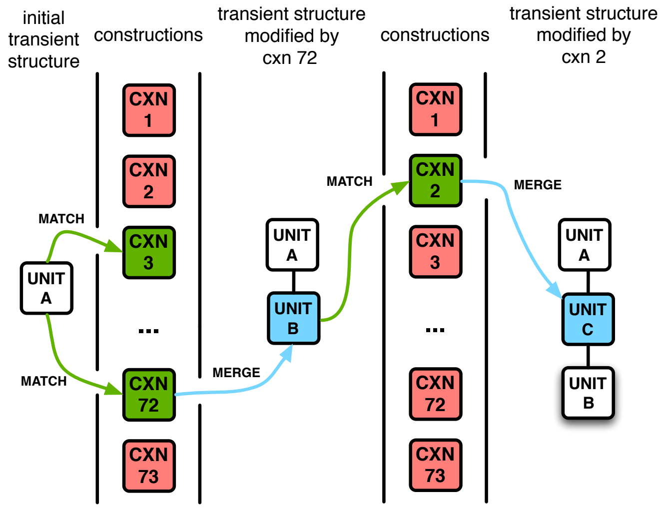

Figure 4.3: This figure shows a schematic view on construction application (Figure adapted from Steels 2011b). Starting from the initial transient structure (left) all constructions in the set of constructions are tested whether they match with the structure. Two constructions match with the initial transient structure. If a construction matches it can merge new information. Construction 72 adds unit C. After the structure has been changed, the process continues and all constructions are checked whether they merge with the transient structure modified by construction 72. Because construction 72 has applied, the transient structure is in a state such that construction 2 can now apply. This was previously not the case. Construction 2 is depending on information provided by construction 72. Subsequently, construction 2 further changes the transient structure and so on and so forth. Often multiple constructions from the set of constructions can apply. For example, construction 3 could also change the initial transient structure. This poses a general problem in processing which is solved by using a search algorithm described later in this section.

4.1.2 Constructions

Constructions are organized in the same way as transient structures. They consist of two poles and the data in each pole are 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 difference 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.

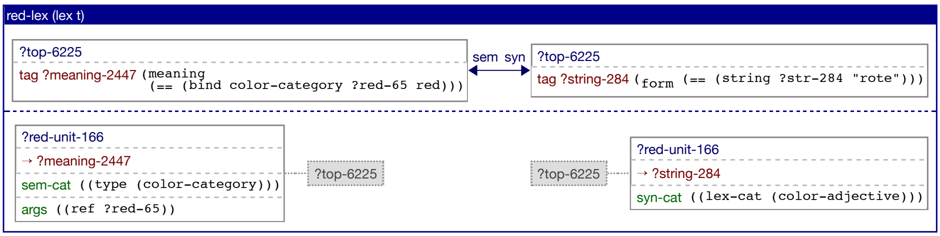

Figure 4.4: Schematic representation of a construction. The two poles of the construction are shown. The top shows the tagged and matching parts of the construction. The bottom shows the hierarchy building part of the construction.

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

(3)
```
((?top-6143
  (tag ?meaning-2381
   (meaning (== (bind color-category ?red-57 red)))))
 ((j ?red-unit-158 ?top-6143)
  ?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 (3), 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:

(4)
```
((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 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.

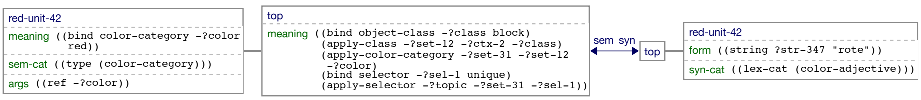

Figure 4.5: Transient structure after the lexical construction applied. The construction has introduced two new units using the J-operator. One on the semantic side and one on the syntactic side. Both units have the same name `red-unit-42`. The construction introduced the string *rote* on the syntactic 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.

4.1.3 Search

Constructions are organized in a pool of constructions. In principle, constructions compete for access to the transient structure in processing. More than one construction can typically apply to the transient structure and the question is how to organize the process if there are multiple constructions that want to change the transient structure. In the absence of a priori rules to prefer one construction over another, each construction that can apply to the transient structure, is tried in a different branch of a heuristically 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*.

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

4.1.4 Design layer

In order to design grammars it has proven beneficial to abstract 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 also for phrasal constructions in other grammars. Templates are defined similar to functions. They have a name and a set of arguments which are specific to the template.

(5)
```
(\emph{template-name} \emph{construction-name}
 \emph{:argument-1} \emph{value-1}
 \emph{:argument-2} \emph{value-2}
 ...
 \emph{:argument-n} \emph{value-n})
```

Let us consider an example. I redefine the lexical construction introduced earlier, using a template called `def-lex-skeleton`.

(6)
```
(def-lex-skeleton red-cxn
 :meaning (== (bind color-category ?cat red))
 :args ((ref ?cat))
 :string ”rote”)
```

If this template is executed it translates into the low-level list representation in (3).

4.2 Open-ended language evolution with FCG

Besides the obvious requirement of computational formalism for linguistic processing for computational experiments, FCG has a number of features that make it an optimal choice for studies in language evolution. FCG is not fixed to a certain set of constructions, a particular grammar layout, a particular set of meanings, or even a particular set of semantic and syntactic categories. FCG solely provides dedicated mechanisms for processing language but makes no actual claims about how a particular phenomenon should be processed in language. This allows different solutions to be explored by grammar designers. But, most importantly, it allows artificial agents to invent different constructions for solving a particular problem in communication, track their success and adapt them until the agents have conventionalized a solution to their particular problem. Language is not a fixed system, but rather a system negotiated by its users to reach communicative goals in a decentralized manner. The fact that there are different solutions to solving the same problem therefore requires formalisms that are designed to be open to change syntactic and semantic categorization, and evolve meaning structure and new constructions. FCG is such a formalism.

From a computational perspective, FCG provides an easily manipulatable data representation, which makes inventing new constructions, changing and adapting semantic and syntactic categories, and introducing hierarchy or movement of data relatively easy. As in 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 the 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 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 research complex language phenomena such as Russian aspect (Gerasymova & Spranger 2010, 2012). The complexity of natural language has without 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 decade. The interested reader is referred to these efforts to get a broader introduction.

Part II

Reconstructing German locative phrases

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. 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 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 structure 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 (1), as adverb as in (2) and as preposition as in (3).

(1) der vordere Block
the.NOM front.ADJ.NOM block.NOM
‘The front block’

(2) der Block vorne
the block.NOM in the front.ADV
‘The front block’

(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). For instance, the following phrase (4) 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.

(4) Stelle den Stuhl vor den Schrank!
Put the chair front.PREP the cupboard
‘Put the chair in front of the cupboard!’

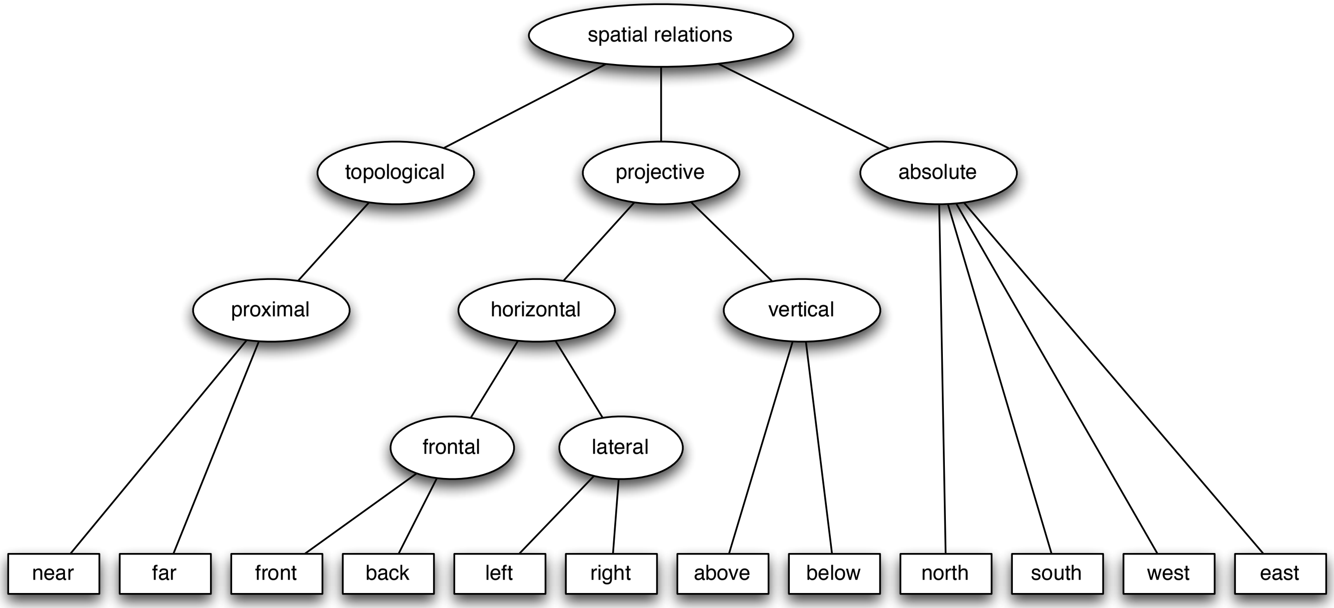

Figure 5.1: Taxonomy of spatial relations discussed in this book. The rectangular items refer to the spatial terms used for denoting the relations. Since in German spatial relations can be expressed in different lexical classes and word forms, English equivalents are used as placeholder.

Another important difference between spatial adjectives and spatial adverbs and prepositions is the potential for *reference objects* or *landmarks*. Example (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 (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 (5)–(6):

(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'

(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, via the spatial term, relates to the figure, in this case *der Block*. 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 (7) an adverb is perspective marked, in (8) a prepositional phrase is perspective marked.

(7) der Block vorne von dir aus
the.NOM block.NOM front.ADV from.PREP the.DAT box.DAT
'The block in front from your perspective'

(8) der Block links der Kiste von dir aus
the.NOM block.NOM left.PREP the.GEN box.GEN from.PREP 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 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 its spatial semantics. This resonates with theories of syntax which put the direct mapping of syntax to semantics at the core of language processing, such as Construction Grammar. 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 I am focusing 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. One needs 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 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 section examine three main questions. First, what is the meaning of phrases such as *der Block links von der Box von dir aus* ('the block left of the box from your perspective', see (8)) and how can we formalize the semantic structure of these utterances? What are the cognitive primitives that are necessary for modeling the semantics of spatial phrases in particular? 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?

6 Spatial semantics

So what is then the meaning of a phrase such as *der Block links von der Kiste von dir aus* ('the block left of the box from your perspective')? Figure 6.1 shows an IRL-network that agents autonomously construct to conceptualize a particular spatial scene. The structure consists of a set of cognitive operations that involve, for example, the construction of regions, the identification of landmarks, the application of perspective transformations and so on. But of course the network also contains references to spatial categories, selectors, and other semantic entities that are processed by cognitive operations. The following sections identify and describe the cognitive operations and the semantic entities underlying locative utterances.

6.1 Representing spatial relations

The semantics of spatial relations is the subject of ongoing debate. The key question is whether there is something like a SEMANTIC CORE of spatial relations, i.e. a core meaning that abstracts from discourse situations as far as possible (Tenbrink 2007), and what that semantic core should be. For many scholars, the semantic core of spatial relations is related to geometric properties (Herskovits 1986; Eschenbach 1999; Tenbrink 2007) in particular prototypical axis (directions) and distances defined using tools like lines, points, vectors, half-planes, etc. (Levinson 1996). Particularly, projective relations, e.g. *vor* ('front') have been studied in this respect and certainly the class of absolute relations, e.g. *nördlich* ('north'), can be conceived in these terms. For instance, Herskovits (1986) describes the meaning of the spatial relation *in front of* as graded concept with the frontal axis as the focal region. This conception links to another important property of spatial language namely its inherent vagueness (Hall & Jones 2008). Many of these proposals are rooted in a strand of psycholinguistic research that is concerned with prototypes and prototypicality effects (Rosch 1978). Here, prototypes or prototypical points in the sensorimotor space are used as representations for spatial categories (for a similar approach to color, see Bleys 2010). For instance, the projective term *links* ('left') can be interpreted for objects which relate to the

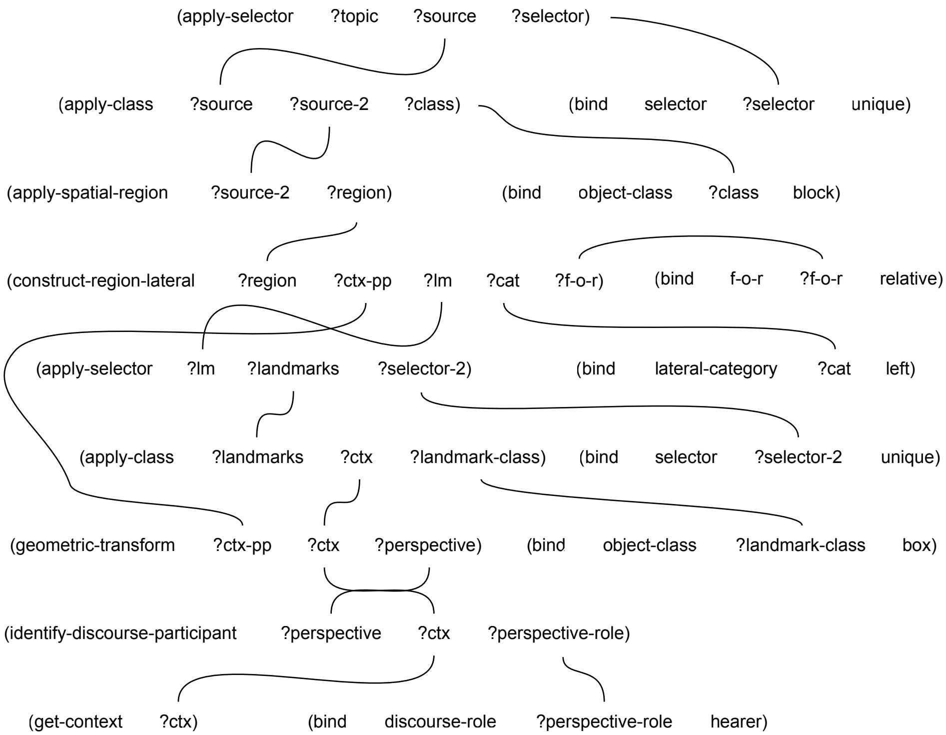

Figure 6.1: An IRL-network representing the meaning of the phrase *der Block links von der Kiste von dir aus* (‘the block left of the box from your perspective’)

reference object by an angle of 90° (prototypical left angle). The more the angle between the target object and the reference system deviates, the less acceptable the spatial relation becomes (Tenbrink 2005a; Herskovits 1986; Gapp 1995).

In this section I focus on the geometric properties of spatial relations and combine them with the prototype approach to spatial categorization. There are two features of the sensorimotor space which are of particular importance for spatial categorization: distance and angle. Two objects always have a certain distance from each other, and if there is a coordinate system available which supports rotation also angles between objects can be measured. Consequently, from a computational point of view, there are two important category types for representing the geometric properties of the German spatial relations discussed in this book `angular-category`, which represent prototypical angles and `proximal-category`, which represent prototypical distances (see Figure 6.2 for an overview).

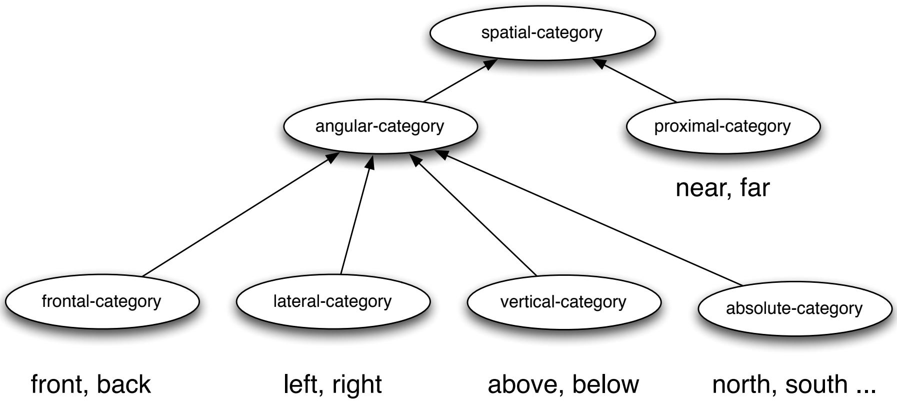

Figure 6.2: Category type hierarchy used in semantic processing

6.1.1 Angular relations

The core semantics of spatial relations is represented using functions that map a particular input location to the applicability degree of a particular category. For prototype based spatial categories degree of applicability amounts to similarity in some spatial dimension, e.g. in the angular dimension for angular relations. Projective and absolute relations are examples of angular categories, with a focal region around the denoted axis. For instance, frontal categories have a high degree of applicability along the frontal axis. Whereas lateral categories have a high degree of applicability along the left and right axis. Similarly, absolute categories have high values of applicability in their respective direction (see Figure 6.1.2 for an overview). In other words, for angular categories, similarity of some location to the category depends on the distance of angles. In order to get a similarity function $sim_a \in [0, 1]$ the angle distance is wrapped in an exponential decay envelope and weighted by a σ which steers the steepness of the exponential decay. High values for σ correspond to a slow decay in similarity the bigger the angular distance, whereas low values for σ correspond to a sharper decline in similarity. Consequently, the following equations defines the degree of applicability given a location l and an angular category c, as the angular distance between c and l,

I think "the bigger the angular distance" should be omitted weighted by σ and run through an exponential decay.

$$\mathrm{sim}_a(l, c) := e^{-\frac{1}{2\sigma_c} d_a(l,c)} \tag{6.1}$$

$$d_a(l, c) := |a_l - a_c| \tag{6.2}$$

In this definition a_l denotes the angle of the position of location l to the coordinate center and a_c denotes the prototypical angle of the category c. Given this definition, one can go ahead and define angular categories, in particular I need to define the prototypical angle for each angular category and the σ. Examples of definitions of angular spatial relations are depicted in Figure 6.1.2.

6.1.2 Proximal relations

Proximal relations relate to some prototypical distance. Two relations are modeled: `near` and `far`. The only difference to the definition of angular categories is that proximal relations rely on the distance channel.

$$\mathrm{sim}_d(l, c) := e^{-\frac{1}{2\sigma_c} d_d(l,c)} \tag{6.3}$$

$$d_d(l, c) := |d_l - d_c| \tag{6.4}$$

In this definition d_l denotes the distance of the location l to the coordinate center and d_c denotes the prototypical distance of the category c.

6.2 Applying spatial relations

Prepositionally and adverbially expressed spatial relations denote spatial regions which are always related to some reference object. For instance, in the phrase *der block links der Kiste* (‘the block to the left of the box’) an object *der Block* (‘the block’) is related to the landmark *die Kiste* (‘the box’) via the spatial category *links* (‘left’). The information about which category is used and which landmark is referred to is packaged in a *spatial region*. In order to represent the difference between the construction of a region as for instance denoted by the prepositional phrase *vor der Kiste* (‘in front of the box’) and the application of that region to filter objects, as in the phrase *der Block vor der Kiste* (‘the block in front of the region’), in other words in order to represent spatial relations, spatial processing is split into two distinct semantic operations. One operation constructs the region and packages the particular landmark and the particular spatial category, the other applies the region as a spatial relation to the objects available in the context.

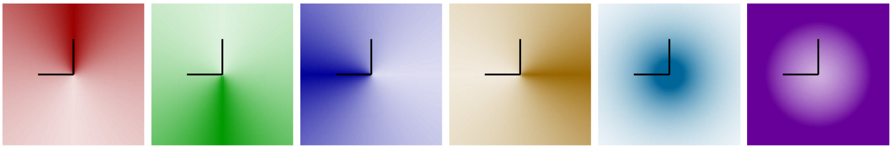

Figure 6.3: Degrees of acceptability shown for prototypical representation of frontal and lateral projective categories. From left to right `front`, `back`, `left`, `right`, `near` and `far` are shown. The opacity of each color denotes acceptability for a particular location in space (x-axes and y-axes each run from -2000.0 to 2000.0). `Front`, for instance, shows a strong acceptability along the x-axes. Definitions of categories are $a_{\texttt{front}} = 0.0$, $a_{\texttt{back}} = \pi$, $a_{\texttt{left}} = \frac{\pi}{2}$ and $a_{\texttt{left}} = -\frac{\pi}{2}$ and $\sigma_{\texttt{front}} = \sigma_{\texttt{back}} = \sigma_{\texttt{left}} = \sigma_{\texttt{right}} = 1.0$. For absolute categories the same definitions exist with `north` being defined like `front` and so forth. The two proximal categories `near` and `far` are defined with $d_{\texttt{near}} = 0.0$, $a_{\texttt{far}} = 2000.0$ and $\sigma_{\texttt{near}} = 700.0$ and $\sigma_{\texttt{far}} = 1200.0$.

6.2.1 Spatial regions and spatial relations

Proximal regions are computed based on the distance prototype of the corresponding category and the landmark. The semantic operation `construct-region-proximal` therefore constructs a specific region based on a spatial category and a landmark. The right image in Figure 6.4 shows an example of a proximal region.

Semantic operation CONSTRUCT-REGION-PROXIMAL

description	Computes a proximal region based on the landmark.
arguments	`?spatial-region` (of type spatial-region) `?source-set` (of type entity-set) `?landmark` (of type point) `?category` (of type proximal-category)

The other operation needed for applying a region is called `apply-spatial-region`. This operation computes the similarity of every object in the context, given a region constructed, for instance by the operation `construct-proximal-region`. For the case of a proximal region, this involves (1) transforming the context so that the landmark is at the center of origin of the coordinate system and (2) applying the similarity function defined in Equation 6.3. This operation in many ways acts as a classifier such as the semantic operations for `color` and `object-class` described in Section 3. For each object in the source set the similarity of this object to the spatial region is computed, based on the constituents of the region, e.g. which category and which landmark was used to define the region. This similarity is combined with the other computed similarities for the object (see Section 3 for description). The operation `apply-region-filter` is general enough to be applicable to all regions, including projective and absolute regions whose description is to follow.[1]

Semantic operation APPLY-SPATIAL-REGION

description	Applies a spatial region, by computing the similarity of the region with every entity in `?source-set`.
arguments	`?target-set` (of type entity-set) `?source-set` (of type entity-set) `?region` (of type spatial-region)

An example of the interplay of the operations `construct-region-proximal` and `apply-region-filter` for a spatial scene can be seen in Figure 6.4. The following table summarizes the similarities computed when applying the region to the context (Figure 6.4 left).

object	distance (mm)	similarity
robot-1	1041.02	0.48
robot-2	1224.53	0.42
box-1	0.00	0.00
obj-266	419.97	0.74
obj-265	938.99	0.51

6.3 Frames of reference

Projective and absolute relations are defined in terms of focal directions. Consequently, for these kinds of relations the rotation of the landmark is an important issue determining the precise applicability of these relationships. For example, what is considered the front direction of a landmark has a direct effect on what is considered a frontal region. It turns out that there is considerable amount

[1] This operation can be very general because its implementation defers different methods of computing similarity for different category types using the method dispatching facility of lisp.

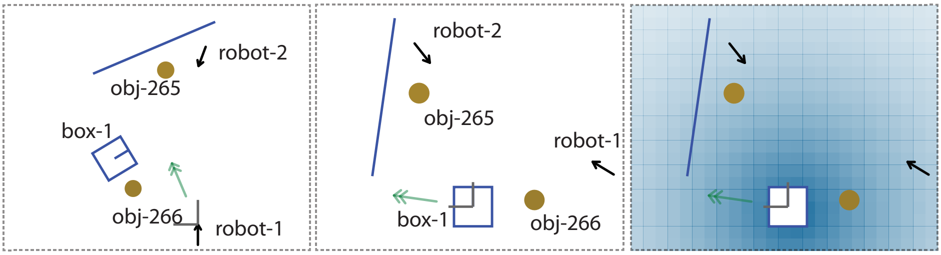

Figure 6.4: When a region is applied, first, the context (left figure) is transformed so that the landmark is at the origin of the coordinate system (middle figure) which is followed by the application of the spatial relation (right figure). Here, these steps are depicted for the category `near` and the landmark `box-1`.

of choice when it comes to how to define the rotation of the landmark. The combination of a coordinate system, in particular its rotation, with a landmark is called REFERENCE SYSTEM. Reference systems have been dealt with in great detail in cognitive semantics and psycholinguistics under the term FRAME OF REFERENCE. Levinson (1996, 2003) identifies three possible frames of reference: INTRINSIC, RELATIVE and ABSOLUTE, all of which denote a particular way of construing a landmark for spatial relationships involving direction. In German all three frames of reference are possible.

Intrinsic frame of reference The intrinsic frame of reference is an object centered coordinate system, meaning that projective categories are applied to the reference object based on particular sides of the object, which are construed as front, back, left and right. Hence, those objects that have something that can be considered as their front (with other sides, identifiable as well, e.g., left, right and back) are eligible to be used as landmarks with an intrinsic frame of reference. Examples of such objects are television sets, where the front is the screen, or houses, where the front is the main entrance or street access, and so forth.

Relative frame of reference The relative frame of reference is a perspective based coordinate system. (See Figure 6.5 for a graphical explanation.) Instead of relying on intrinsic features of the reference object for determining the particular layout of the coordinate system, the rotation of the coordinate system is determined by its angle to an explicitly or implicitly given perspective. Hence, the front of an object is induced by the particular perspective on the scene. For example, *vor dem Baum* ('in front of the tree') implicitly refers to a perspective, because trees do not have an intrinsically determined front, and it is the position of the observer together with the position of the tree that designates the precise region denoted as front.

Absolute frame of reference Absolute frames of reference construe the landmark using an external rotation. Neither intrinsic properties nor the perspective on the landmark determine the layout of the coordinate system, but rather geocentric features of the environment, for instance cardinal directions as in *nördlich* ('north') or the direction of gravity as in *über* ('above') govern what the precise layout of the reference system is.

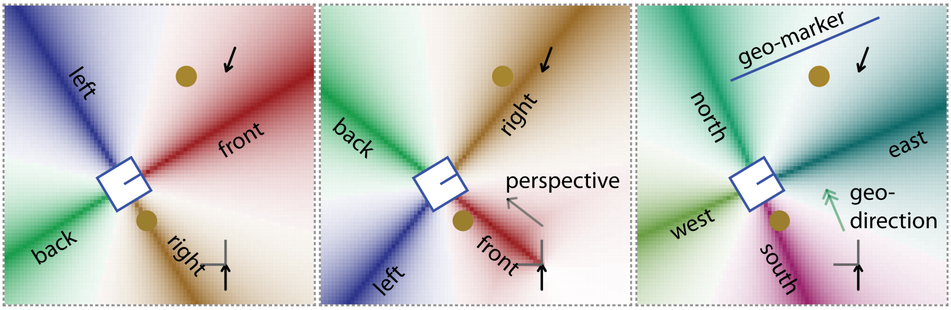

Figure 6.5: Frames of reference have a profound effect on spatial relations. In all three pictures spatial relations are shown with `box-1` as reference object. The left figure shows the landmark construed with an intrinsic frame of reference which orients the coordinate system such that the front of the landmark (little blue line in the landmark) corresponds to the frontal spatial relation with left, right and back projective relations aligned accordingly. The middle figure shows the same landmark, but with a relative frame of reference constructed from the perspective of robot `robot-1` which entails that frontness corresponds to a region between the perspective and the landmark. Left and right in relative frames of reference are aligned not with respect to front, but rather are parallel to the left and right side of the perspective `robot-1`. The right figure shows an absolute frame of reference applied to the landmark `box-1`. Cardinal directions are aligned based on a geocentric direction induced on the scene by a marker.

The differences in processing frames of reference are captured using distinct semantic operations. In other words, absolute relationships, such as *nördlich* ('north') or *über* ('above') demand other semantic operations as relative and intrinsic ones. For instance, processing of phrases involving absolute regions, e.g. *nördlich* ('north'), is represented using the `construct-region-absolute` operation, which takes a landmark and transforms the context with respect to the landmark, subsequently applying a rotation that follows from a geocentric direction. Some of the spatial scenes recorded by the robots feature a geocentric marker on the wall. The direction towards this marker defines the direction to the north. Figure 6.5 gives an example.

Semantic operation CONSTRUCT-REGION-ABSOLUTE

description	Computes an absolute region based on the landmark and the absolute frame of reference which must be available in source set.
arguments	`?spatial-region` (of type spatial-region) `?source-set` (of type entity-set) `?landmark` (of type point) `?category` (of type absolute-category)

Frontal prepositions, e.g., *vor* ('front') and *hinter* ('back'), can have both intrinsic and relative readings. Both readings are incorporated into a single operation, which construes the landmark both in relative and intrinsic way signified by an additional parameter of type `f-o-r` (frame of reference) to the operation. For the relative reading the perspective on the scene additionally influences the layout of the region. The viewpoint on the scene constrains front regions in such a way that only those locations which are between the perspective and the landmark have a high degree of applicability.

Semantic operation CONSTRUCT-REGION-FRONTAL

description	Computes a frontal region based on the landmark and the relative or intrinsic frame of reference.
arguments	`?spatial-region` (of type spatial-region) `?source-set` (of type entity-set) `?landmark` (of type point) `?f-o-r` (of type f-o-r) `?category` (of type frontal-category)

Vertical relations can be construed with an absolute, relative or intrinsic frame of reference. In the case of absolute frames of reference the orientation is derived from gravity. For the purpose of this book only absolute frames of reference readings are implemented in the operation `construct-region-vertical`.

6.4 Internal and external regions

Another line of processing distinctions can be made between internal and external relations. *North* and *south*, but also prepositional use of *front*, e.g. *vor*, are referring to external regions, that is the region lies outside of the landmark. Internal regions on the other hand lie inside the reference object. Adverbs such as *vorne* ('front') denote such regions inside the landmark. Consequently, they require separate treatment and there is a specific operation for handling the internal processing of frontal relations `construct-region-frontal-internal`.

Semantic operation CONSTRUCT-REGION-FRONTAL-INTERNAL

description	Computes an internal frontal region based on the landmark and the relative or intrinsic frame of reference.
arguments	`?spatial-region` (of type spatial-region) `?source-set` (of type entity-set) `?landmark` (of type point) `?f-o-r` (of type f-o-r) `?category` (of type frontal-category)

Internal lateral regions are not clearly marked in syntax. While lateral prepositions clearly denote an external region, lateral adverbs can be used more varied which becomes apparent since they can be complemented both by *von* headed prepositonal phrases, as well as *in* headed prepositional phrases. In other words, the interpretation of an adverb depends on the complement. If there is no complement both readings internal and external are possible.

Semantic operation CONSTRUCT-REGION-LATERAL

description	Computes an internal frontal region based on the landmark and the relative or intrinsic frame of reference.

arguments `?spatial-region` (of type spatial-region)
`?source-set` (of type entity-set)
`?landmark` (of type point)
`?f-o-r` (of type f-o-r)
`?category` (of type frontal-category)
`?region-layout` (of type region-layout)

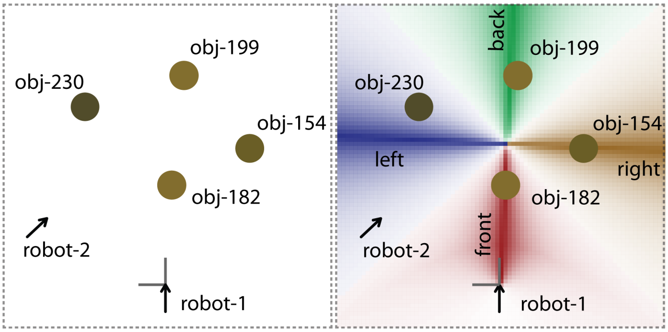

Figure 6.6: Group-based reference requires computing a landmark based on a group of objects. The centroid of the group of objects, in this case all blocks in the context, is construed using a relative frame of reference. Consequently, object `obj-182` could be described in German as *der vordere Block* (‘the front block’).

6.5 Group-based reference

The semantics of German spatial adjectives is best understood in terms of group-based reference.[2] In order for a group of objects to function as a reference object, the group needs to be construed as a landmark object, and in particular its position needs to be established. One way of computing a position for a group of objects, is to use the spatial centroid (center of mass) of the group as the position of the reference object. Additionally, a frame of reference, in other words a rotation, needs to be chosen. This choice depends on the spatial relation. For absolute relations the frame of reference is determined by an absolute frame of reference. In the absence of intrinsic features of the group of objects, a relative frame of reference seems to be the natural choice, as a spatial group of objects has

[2] Recent evidence (Moratz & Tenbrink 2006) points to more variety in interpretation. For the purpose of this book, however, I choose to model group-based semantics only.

no intrinsic front without including further constraints. Together, the centroid and the respective frame of reference sufficiently describe the reference system in order for spatial relations to be applied.

Semantic operation APPLY-SPATIAL-CATEGORY-GROUP-BASED

description	Applies the spatial category based on a group-based landmark.
arguments	`?target-set` (of type entity-set) `?source-set` (of type entity-set) `?category` (of type spatial-category)

Group-based reference offers a technical challenge of how to compute the group used for reference in the first place. Given that all objects in the context are scored by successive operations, how can a group of objects be established, in order to compute the spatial centroid? For instance, in phrases like *der linke block* ('the left block'), the spatial adjective is part of a noun phrase, which primarily denotes the type of objects that constitute a set, namely the set of all block which relates to the group that the reference should be based on. The implementation of the semantic operation `apply-spatial-category-group-based` relies on a well-known clustering algorithm called *k-means* (Lloyd 1982) in order to find the group of objects in the input set. The scores of all objects in the input set are used to divide the input set into two groups and all elements in the group with the higher score centroid are used to compute the spatial landmark.

6.6 Perspective marking

The perspective of relative frames of reference is sometimes explicitly marked by speakers. In the example *der Block links von der Kiste von dir aus* ('the block left of the box from your perspective'), the speaker choose to provide a perspective *von dir aus* ('from your perspective'). Only relative interpretations of a spatial phrase can be perspective marked. The perspective itself marks the position of the viewpoint, and therefore also the view direction, i.e. rotation.

Intrinsic frames of reference computationally behave very similar to perspective-marked relative frames. Some argue therefore, that an intrinsic reference system is in essence a conflated relative reference system that coincides with the perspective on the scene (Levinson 1996). Hence, intrinsic reference systems are never perspective marked, since they already include position and orientation. However, perspective marking is only compatible with relative frames of reference

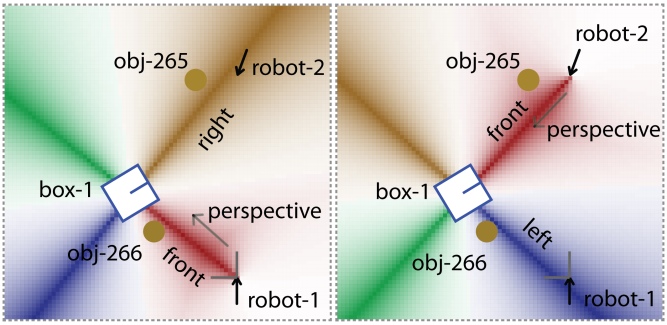

Figure 6.7: The two figures show the influence perspective has on relative frames of reference. The left figure shows the landmark `box-1` when construed with a relative frame of reference from the perspective of `robot-1`. In this configuration object `obj-266` is in front and object `obj-265` is to the right. Whereas when the same landmark `box-1` is construed with a relative frame of reference from the perspective of `robot-2` it is object `obj-265` which is in front, whereas object `obj-266` is to the left.

and excludes intrinsic the usage of intrinsic or absolute frames of reference. On the other hand, relative reference systems always explicitly or implicitly mark perspective, since they cannot be conceived without since by definition relative reference systems always construe the world from a perspective (see Figure 6.7).

Perspective on a scene is changed by an operation that transforms a complete spatial context as if it had been perceived from a particular viewpoint.

Semantic operation GEOMETRIC-TRANSFORM

description	Transforms the context to be viewed from a particular perspective. Notice that perspectives require both a particular point of view but also a rotation.
arguments	`?transformed-context` (of type sensory-context) `?context` (of type sensory-context) `?perspective` (of type pose)

6.7 Discussion

6.7.1 Functional constraints

Purely geometric accounts of spatial semantics have been criticized on the basis of psycholinguistic studies that reveal for many spatial relations, that additional functional constraints influence their applicability. Studies in particular for topological relations, including *in* and *on*, but also for projective relations such as *over* and *under* (Coventry et al. 2001) as well as *in front* and *behind* (Carlson-Radvansky & Radvansky 1996) have led to new proposals (Coventry et al. 2005) as to how to include functional considerations into the semantics of spatial terms (see also Coventry & Garrod (2004) for an overview). For instance, whether or not an umbrella is *over* a person depends on the direction of rain which can come from different angles. I do not account for functional constraints for two reasons: (1) because it requires detailed functional models of objects which as of now are rarely available in robots and (2) the current model theoretically can incorporate such models once they become available. In order to acquire functional knowledge, such as that chairs are for sitting, tables are used to put things on and so forth, robots need to interact robustly and repeatedly with objects of this kind, in particular using complex interactions in which objects take on functional roles. Many of these skills, e.g. basic actions such as sitting, are contemporary research fields in robotics and still need to see significant progress before they are generally available. On the other hand, functional approaches to semantics are typically committed to conceive the application of spatial relations in terms of degree of applicability. In other words, functional models are essentially mappings of locations and landmarks to some number representing the degree in which some relation is deemed acceptable. This is precisely the basis of the semantics advocated in this chapter. Semantic operations compute degrees of applicability. Consequently, once a functional model can be established in terms of similarity, it can readily be incorporated into the current model by exchanging semantic operations.

6.7.2 Contextual factors

Besides functional constraints, contextual factors affect the applicability of spatial terms. For instance, for proximal relations: *prototypical size* (Gapp 1994), but also *object salience* (Regier & Carlson 2001), and *object interference* (Kelleher & Costello 2009) seem to play a role. Just to give an example, the prototypical size factor can explain why the proximal region *nahe des Gebäudes* (‘near the building') is larger than that of *nahe dem Apfel* ('near the apple') (example adapted from Gapp 1994) by the difference in typical size of buildings and apples. Constraints such as the prototypical size, as well as the influence of object salience, are easily integrated into the model, but just as for functional constraints do not affect the basic assumptions of the model. The third constraint – the object interference constraint – refers to the interference by other objects that are for instance closer than the related object. Such constraints are better treated under the problem of how to choose spatial relations which is inevitably connected to the particular communicative goal. For instance, in a discrimination task other categories might be more relevant for the task than in object location description tasks. Such processes are dealt with in detail in Chapter **??**.

6.7.3 Other modeling approaches

Spatial semantics is an important and vibrant research area. Many different proposals are currently being made. Some suggest the use of formal ontology engineering (Bateman et al. 2007; Bateman 2010; Bateman et al. 2010) as a tool for enhancing spatial language interpretation by artificial systems. Others suggest to use formal reasoning techniques and representations (Freksa 1991; Cohn & Hazarika 2001). The system presented in this book can benefit from these extensive approaches in the sense that the detailed modeling of spatial representation and spatial reasoning could enhance our modeling approach. On the other hand, in this book modeling serves the goal of establishing basic concepts, e.g. spatial relations, so that we can later study their evolution. This is the reason why more elaborate modeling approaches are avoided. Engineering robust and extensive solutions for the processing of spatial language is a valid goal in itself, but it is only one aspect of this book.

6.7.4 Summary

This chapter gave an account of the semantic core of German spatial relations in terms of geometric constraints, frames of reference and perspective. Spatial relations have been defined as graded categories, whose application is governed by semantic operations. It remains to be shown (1) how conceptualizations of a scene given a concrete communicative goal are achieved (a problem that can be summarized in how and which spatial relations should be chosen), (2) how spatial categorization fits into larger semantic structures for spatial phrases such as *der Block links von der Kiste* and (3) how these semantics interact with language in production and parsing of spatial phrases.

7 Syntactic processing

The syntax of German locative phrases mirrors the complexity of spatial semantics. The main task of syntactic processing is to allow agents to express themselves by translating semantic structure into proper German syntax and back. This chapter reports on the syntactic processing of German locative phrases using Fluid Construction Grammar. It provides an overview of the mapping from semantics to syntax and zooms in on different aspects of the implementation for dealing with complicated syntactic phenomena such as the German case system.

Syntactic processing of German spatial language is primarily a problem of orchestrating intertwined information processing. Great care has to be taken to ensure that the ordered application of constructions in production and parsing of spatial utterances can proceed efficiently,without excessive branching in search. In many cases information needed for processing is spread in the semantic structure to be produced or in the utterance to be parsed. For instance, lexical constructions might able to decide on word stems based on semantic entities in the semantic structure, but already the decision which lexical class to use for expressing some lexical item in production requires a larger broader of the semantic structure to be produced. Even more so, in order to decide on the actual word form including German morphology, a whole array of syntactic information is to be considered. For instance, case, gender and number marking of spatial adjectives in prepositional noun phrases requires collection of information from the noun about its grammatical gender, and from the preposition about the required case. This chapter shows how advanced techniques in FCG can be applied to organize efficient processing while allowing grammar designers to build extendable and concise grammars.

This chapter starts by presenting the general ideas behind the design of the grammar (Section 7.1) and identifying core issues that have to be resolved in order to arrive at an operational system. The remaining Sections 7.2–7.5 show how to deal with these issues.

7.1 Overview of syntactic processing

One of the main problems of syntactic processing in a system that is integrated with procedural semantics such as IRL is the problem of how semantic structure such as IRL-networks are expressed in language. The following are the two main ideas originally presented in Steels & Bleys (2005).

Lexicon Lexical constructions directly map semantic entities to content words.

Grammar Grammatical constructions broadly speaking encode the cognitive operations and the links between them and semantic entities. Consequently, grammar provides information on how semantic structure is combined (Steels 2011a) by modulating the syntactic expression of the lexical items through grammatical markers, word order, morphology, etc.

German requires fine grained distinctions of constructions in order to facilitate and coordinate processing. The grammar is organized into certain types of constructions each providing different information (see Steels 2011a for the original idea).

Lexical Constructions These constructions map semantic entities to stems and back. Hand in hand with morphological constructions they are responsible for the expression of lexical items. Lexical constructions introduce lexical units in the transient structure that are used to assemble information necessary for decisions on the lexical class and the word form of the semantic entity in production. In parsing these units gather information required for the semantic interpretation of a semantic entity in the surrounding semantic structure. Lexical constructions introduce abstractions into the transient structure that allow to go from concrete semantic entities to the class they are in. For instance, *links* ('left') is a lateral spatial category such as *rechts* ('right'). Both behave similar in semantics as well as in syntax. Lexical units introduce such type information to allow hierarchy in processing to take advantage of this information.

Functional Constructions This class of constructions is related to how a particular semantic entity is used in processing. For instance, in parsing, when a spatial relation is observed as an adjective, this licenses the introduction of the cognitive operation `apply-region-group-based` which represents how the spatial category is supposed to be applied. Consequently, functional constructions handle and process lexical classes and map them to particular cognitive operations. These constructions introduce functional units into the transient structure, which assemble all information related to lexical classes. On the semantic side these units gather information related to the output and input of the cognitive operation. This information is particularly used by phrasal constructions to combine functional units into larger phrases. There are functional constructions for determiners, nouns, spatial adjective, spatial adverbs and prepositions.

Phrasal Constructions These constructions, as the name suggests, are organizing both larger syntactic structure, i.e. phrases, and larger semantic structure by linking functional units. For instance in parsing, when observing a spatial adjective and a noun in the correct German word order, the adjective-noun-phrase construction links the processing of these items on the semantic side. In turn, the new constituent can be further combined with determiners which happens when the determined-noun-phrase construction applies. Phrasal construction do the work of processing grammatical relations between constituents.

Morphological Constructions Certain constructions deal exclusively with morphology. They are responsible for determining and processing word forms, i.e. the string, for expressing lexical items. In German the concrete form of a lexical item is often determined by the larger syntactic structure. For instance, a spatial adjective such as *links* ('left') in a determined adjective noun phrase has to agree with the gender and the number of the noun and with the case of the determined noun phrase. This requires that in production the form of a lexical item is determined when this information is available, which requires that constructions such as the determined-noun-phrase construction have supplied this information. In parsing, morphological constructions function as word form recognizers.

Semantic Constructions Lastly, there is a special type of constructions which is only used in parsing for handling semantic ambiguity. They are discussed in detail in Section **??**.

Figure 7.1 shows the sequence of construction application in production and in parsing. In processing, the constructions are part of a large pool of constructions, which are applied based on which construction can apply at a specific point in time. Syntactic processing proceeds bottom-up which means from the lexical items to the phrasal level. For instance, phrasal constructions require functional units to function which in turn depend on lexical constructions. In every step information is assembled in hierarchical units, e.g. lexical units, which expose

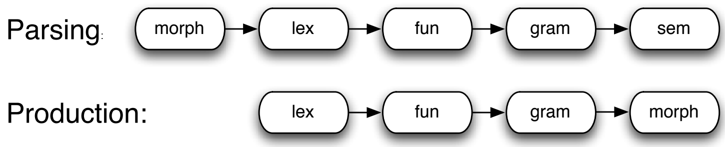

Figure 7.1: Construction application in the German spatial language grammar discussed in this chapter. In parsing, morphological constructions apply first followed by lexical and grammatical constructions. Finally, the semantic constructions important for handling semantic ambiguity apply. In production, these constructions are not applied. In contrast to parsing, morphological constructions apply in production at the very end in order to decide on the actual form used in the utterance.

new information, in order for constructions to use this information. This process can also be understood as a process of gradual categorization and assembly of constituents into larger structure both on the syntactic and on the semantic side. For instance, lexical constructions abstract from the concrete semantic entity such as the spatial category `left` and provide information about the semantic type so that functional constructions can use this information and apply to groups of spatial relations.

The following three sections look in more detail at the implementation of different constructions such as lexical and functional (Section 7.2) phrasal constructions for landmarks and landmark complements (Section 7.3) and high-level phrasal constructions (Section 7.4). The chapter concludes by discussing how to deal with the German case system (Section 7.5).

7.2 Lexical classes – lexical and functional constructions

In German the same spatial relation can be expressed using different lexical classes (adjective, adverb or preposition). The following shows an example of the spatial relation `front` expressed as adjective (Example 1), adverb (Example 2), and prepositions (Example 3) which are repeated here for convenience from Chapter 5.

(1) der vordere Block
the.NOM front.ADJ.NOM block.NOM
'The front block'

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

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

Lexical classes are an example of a many-to-many mappings. Every German spatial relation discussed in this book can be expressed in different lexical classes. Vice versa, every lexical class has a number of relations that are part of it. For instance, the spatial relations `front`, `back`, `left`, can all be expressed as adverb. However, not all spatial relations can be expressed as adverbs. The proximal relations *nah* ('near') and *fern* ('far') do not occur in adverbial form.[1] Table 7.2 gives an overview of lexical classes and associated forms. The table shows that different groups of spatial relations partake in different lexical classes. For instance, all spatial relations except for proximal relations such as `near` and `far` can be expressed as adjectives. Projective relations such as `up`, `front` and `left` can be expressed as adverbs. Only vertical and lateral relations can also be genitive prepositions, whereas frontal relations can only be dative prepositions. Vertical relations can be both genitive and dative prepositions.

The problem of choosing a lexical class in production and finding the lexical class in interpretation is solved by a careful setup of the interaction of functional and lexical constructions. I use a particular design pattern called ACTUAL-POTENTIAL.[2] The design pattern allows to store possible lexical classes in the form of potentials for each lexical item directly in the lexical units of the transient structure. Subsequent functional constructions can constrain their application based on the potential of the lexical item, consequently, ruling out non standard

[1] Linguistic analysis is made difficult by diverging vocabulary in linguistics and different usages of the same term by different schools. I use the term adverb here for spatial relations such as *vorne* ('front') that can be followed by prepositional phrases and used as postmodifiers on determined noun phrases e.g. *der Block vorne in der Kiste* ('in the front area of the box').

check translation

[2] The pattern is inspired by earlier work on argument realization (van Trijp 2008) which is also a many-to-many mapping problem.

Table 7.1: Lexical Classes and word forms for German spatial relations (adapted in part from Tenbrink 2007). This by no means is an exhaustive list of spatial relations, lexical classes or lexical forms in German spatial language, but it is the part of German relevant for this book. Items marked with * seem to be possible, but due to being unconfirmed in the reviewed literature are omitted.

Relation	Adjective	Adverb	Preposition [POSS]	Preposition [DAT]
`up`	*ober*	*oben*	*oberhalb*	*über*
`down`	*unter*	*unten*	*unterhalb*	*unter*
`front`	*vorder*	*vorne*	–	*vor*
`back`	*hinter*	*hinten*	–	*hinter*
`right`	*recht*	*rechts*	*rechts*	–
`left`	*link*	*links*	*links*	–
`near`	*	–	–	*nahe*
`far`	*	–	–	*fern*
`north`	*nördlich*	*	*	*nördlich*
`south`	*südlich*	*	*	*südlich*
`west`	*westlich*	*	*	*westlich*
`east`	*östlich*	*	*	*östlich*

usage, e.g., expressing proximal relations as adjectives. The same is used on the semantic side, where the semantic type hierarchy of the semantic entity is stored as a list of potentials.

The actual-potential technique allows to distribute decision making across lexical and functional constructions by separating the specification of options (potentials) from the *actual* decision. Possible choices are explicitly stored in the form of disjunctive potentials in the transient structure thereby signaling to subsequent constructions which choices are possible which allows subsequent constructions to constrain their application by observing potentials and triggering only when the right potentials are present. Before we jump to the application of the actual-potential design pattern, we need to consider the lexical and functional constructions that are involved.

The fact that Examples (1)–(3) refer to the same projective category `front` is expressed by the lexical construction for the spatial relation `front`. The construction maps the reference to the spatial relation onto the word stem *vor*. The following template shows the lexical construction for the category `front`:

(4)
```
(def-lex-cxn
 (def-lex-skeleton front-cxn
  :meaning (== (bind frontal-category ?cat front))
  :args ((ref ?cat))
  :stem ”vor”))
```

Lexical constructions capture the similarity of different syntactic and semantic usage scenarios of the same category. They encode that no matter how the lexical item is used in the larger semantic and syntactic structure it refers to the same semantic entity, e.g. spatial relation.

Functional constructions map a particular lexical class to syntactic and semantic properties thereby elevating lexical items to constituents in grammatical structure. On the semantic side, functional constructions trigger on semantic operations used in conceptualization. On the syntactic side, the constructions provide syntactic functions and syntactic classes in order for grammatical constructions to be able to build grammatical structure out of functional units. Below are the skeletons for the functional constructions of spatial adjective, frontal adverbs and frontal prepositions:

(5)
```
(def-fun-cxn spatial-adjective
 :meaning (== (apply-spatial-category-group-based
              ?target ?source ?category))
 :args ((ref ?target)(src ?source)(cat ?category))
 :sem-function (modifier)
 :syn-function (adjectival))
```

(6)
```
(def-fun-cxn frontal-adverb
 (def-fun-skeleton frontal-adverb
  :meaning (== (construct-region-frontal-internal
           ?target ?source ?landmark ?category ?f-o-r))
  :args ((ref ?target)(src ?source)
           (cat ?category)(landmark ?landmark))
  :sem-function (modifier)
  :sem-class (region internal-region relative-region)
  :syn-function (adverbial)
  :syn-class (adverb)))
```

(7)
```
(def-fun-cxn frontal-preposition
 (def-fun-skeleton frontal-preposition
```

```
:meaning (== (construct-region-frontal
           ?target ?source ?landmark ?category ?f-o-r))
:args ((ref ?target)(src ?source)
         (cat ?category)(landmark ?landmark))
:sem-class (angular-relationship)
:syn-class (angular-preposition)))
```

These constructions introduce constructional meaning, e.g. `construct-region-frontal`, together with semantic and syntactic potentials. One is the functional role of the unit in the larger syntactic structure, here denoted by `syn-function`. Aside from these two constructions, there are a number of other important functional constructions. In particular the difference in semantics of lateral and frontal adverbs, but also between projective, topological and absolute relations are each captured by separate functional constructions (see Table 7.2 for an overview).

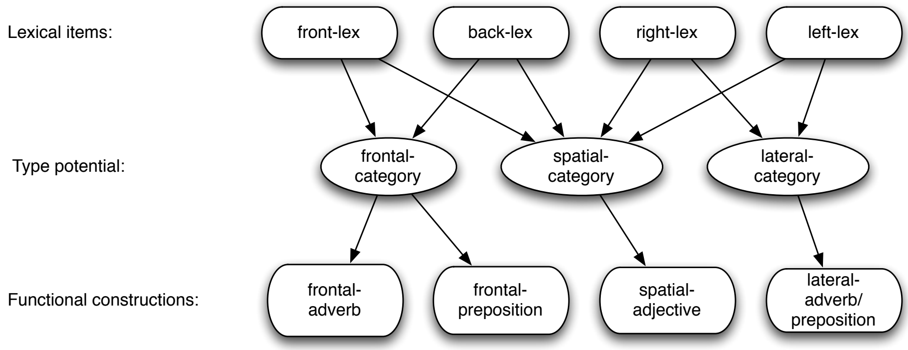

Figure 7.2: Subset of the mapping of lexical items to functional constructions.

7.2.1 Encoding type and lexical class potentials

In order to solve the many-to-many mapping problem, lexical and functional constructions are extended using the *actual-potential* design pattern on the syntactic and semantic pole of constructions. Lexical constructions merge semantic and syntactic potentials into lexical units. This information is used by functional constructions to constrain their application. On the semantic side, the `type` constraints are rooted in the type hierarchy of spatial relations, whereas on the syntactic side, the lexical class lex-cat potential encodes which functional constructions can apply. The constraints take the form of disjunctive lists of potentials.

Table 7.2: Overview of a subset of the functional constructions of the German grammar. The three columns `lex-type`, `lex-cat` and `operation` show the requirements for each construction. The four columns `sem-functions`, `sem-class`, `syn-functions` and `syn-classes` detail the syntactic and semantic functions and classes that the constructions introduce.

name	lex-type	lex-cat	operation	sem-functions	sem-classes	syn-functions	syn-classes	examples
spatial-adjective-cat	spatial-category	spatial-adjective	apply-spatial-category-group-base	modifier	–	adjectival	–	*linke*, *rechte*, *vordere*, *hintere*
frontal-adverb	frontal-category	frontal-adverb	construct-region-frontal	modifier	relative-region, frontal-region, region	adverbial	frontal-adverb, adverb	*vorne*, *hinten*
frontal-preposition	frontal-category	frontal-preposition	construct-region-internal	–	angular-relationship	–	angular-preposition	*vor*, *hinter*
spatial-preposition-cat	spatial-category	spatial-preposition	construct-region-proximal	relationship	–	preposition	–	*an*
lateral-adverb–preposition	lateral-category	lateral-adverb–preposition	construct-region-lateral	modifier	relative-region, angular-relationship, …	adverbial	lateral-adverb, angular-preposition, …	*links*, *rechts*
noun-cat	object-class	noun	apply-object-class	identifier	–	nominal	–	*Block*, *Kiste*
article-cat	selector	article	apply-selector	determiner	–	determiner	–	*der*, *die*, *das*

For example, since all angular categories, i.e. projective and absolute categories, can be used as adjectives, the lexical constructions for these categories feature the `type` potential `angular-spatial-category`, as well as the syntactic `lex-cat` potential `spatial-adjective`. The adjective construction constrains itself to only apply to lexical units that have this potential thereby licensing the application of the spatial adjective construction. The lexical units for proximal relations, such as `near` and `far` do not have these potentials, hence, the adjective construction cannot apply. Other fine-grained distinctions can be modeled as well. Lateral and frontal projective categories differ in how their corresponding adverbs behave syntactically and semantically which necessitates two functional constructions, one for lateral adverbs and one for frontal adverbs. Consequently, the potentials in frontal category units differ from lateral ones. They feature the type `frontal-category`, where lateral lexical constructions (i.e. for `left` and `right`) provide the `type` potential `lateral-category` (see Figure 7.2 for `type` potentials of projective lexical constructions and Figure 7.3 for `lex-cat` potentials)

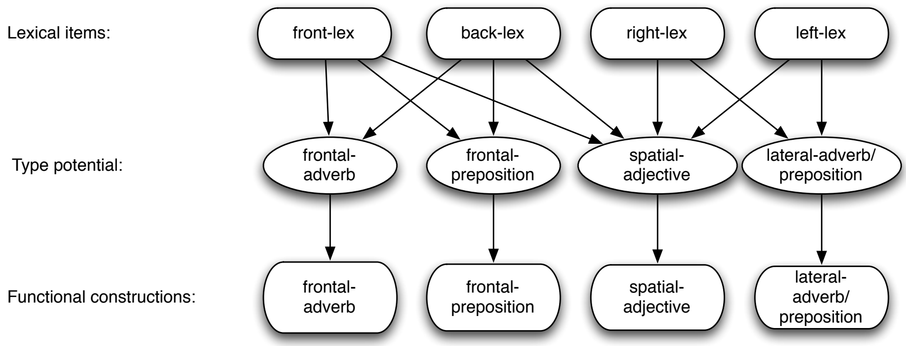

Figure 7.3: Subset of the mapping from lexical items to functional constructions.

7.2.2 Technical realization

The actual-potential design pattern is easily implemented using a dedicated template for extending lexical constructions. Example (4), for instance, is supplemented by the following two templates which introduce potentials on the semantic and syntactic side.

(8)
```
(def-add-potential front sem sem-cat type
  (angular-spatial-category
```

```
  projective-category
  frontal-category))
(def-add-potential front syn syn-cat lex-cat
 (spatial-adjective frontal-adverb frontal-preposition))
```

These two templates specify the `type` and `lex-cat` potentials and directly translate into attributes in the `front` lexical construction:

(9)
```
(...
 (J ?front-unit ?top ()
  ...
  (sem-cat
   (==
    (type
     ((actual ?type-value)
      (potential
       (angular-spatial-category
        projective-category
        frontal-category)))))
   ...)
...)
 <-->
(...
 (J ?front-unit ?top ()
  ...
  (syn-cat
   (==
    (lex-cat
     ((actual ?lex-cat-value)}
      (potential
       (spatial-adjective
        frontal-adverb
        frontal-preposition))))
    ...))
...)
```

Importantly, the template `def-add-potential` not only adds the `potential` attribute but also an attribute called `actual`. This attribute, as we will see in the next paragraphs, is automatically set to a variable in the lexical construction and is used to store which `type` attribute is used. If one of the potentials is picked up, for instance by a functional construction, the `actual` attribute is also set.

The information stored by the lexical construction in the transient structure allows functional constructions to choose the potential in which they are interested and to constrain their own application. This process can be seen in an extended version of the functional spatial adjective construction:

(10)
```
(def-require-potential spatial-adjective
  ?cat-unit sem sem-cat
 type angular-spatial-category)
(def-require-potential spatial-adjective
 ?cat-unit syn syn-cat
 lex-cat spatial-adjective)
```

In order for the spatial adjective construction to apply, these templates express that certain potentials need to be present in the transient structure. More precisely, the `type` potential `angular-spatial-category` and the `lex-cat` potential `spatial-adjective` need to be there.

The template for spatial adjectives translates into the following feature structure (for illustrative purposes, only the semantic side is shown here):

(11)
```
(...
 (?cat-unit
  (sem-cat
   (==
    (type
     ((actual angular-spatial-category)
      (potential
       (==! angular-spatial-category))))
    ...))
...)
```

This construction can only apply if the `type` potential of the lexical constituent in the transient structure imperatively includes `angular-spatial-category`. Additionally, it requires the `actual` attribute to be `angular-spatial-category` or a variable. There are two things to note here: the use of the `==!` operator for potentials and the handling of the `actual` attribute. The `==!` operator only unifies and never merges, which means that neither in production nor parsing can a missing potential be merged. The specified potential always has to be present, in this case on the semantic side, but for the `lex-cat` potentials, the case is vice versa on the syntactic side. Consequently, choosing a potential does not change the potential in the transient structure. The second feature, the `actual` attribute, must be equal to `angular-spatial-category` or a variable in order for the spatial adjective construction to apply. If the attribute is a variable, then that variable is bound to `angular-spatial-category`, and hence the application of the spatial adjective construction modifies the transient structure and sets the `value` attribute to the required potential. Of course, the corresponding potential also has to be present for the construction to apply in the first place (see Figure 7.4)

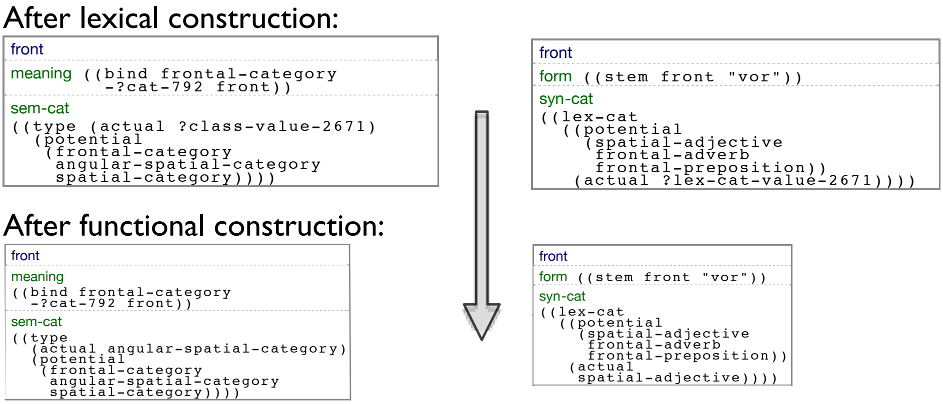

Figure 7.4: Interaction of lexical constructions with functional constructions in production of *vordere* ('front'). The arrow signifies the order of application. Left, the `vordere` unit on the semantic side of the processed transient structure is shown. Right, the syntactic unit is shown. The transient structure actually contains more units, and the units themselves contain more features, but everything has been shortened for illustrative purposes. The top row shows the lexical unit after the application of lexical constructions, which have equipped the lexical unit with potentials for `type` on the semantic side, and `lex-cat` on the syntactic side. Both of these potentials have no value assigned to them yet. It is only after the application of the functional construction of spatial adjective that both have values assigned to them, `spatial-category` for `type` and `spatial-adjective` for `lex-cat`.

This split into `value` and `potential` and the ability to interact via these two attributes is not only interesting for grammar designers who can track the application of constructions by tracing the actually chosen potential, but it also plays an active role in processing. In parsing, the lexical class of a word is decided by morphological constructions. The morphological constructions apply first when parsing an utterance and they provide a value for the `actual` attribute. For instance, when observing the form *vorne*, the morphological construction responsible for the string *vorne* triggers and adds the information to the transient structure, namely that an adverb was observed in parsing (see Figure 7.5 for a schematic overview).

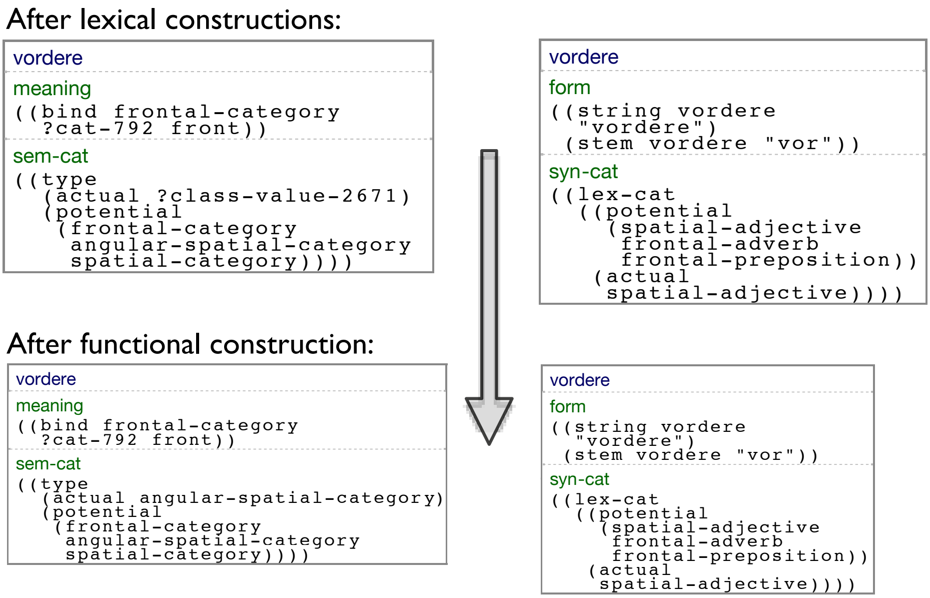

Figure 7.5: Interaction of lexical constructions with functional constructions in parsing *vordere* ('front'). Lexical constructions apply before functional constructions. The `vordere` unit on the semantic side of the processed transient structure is shown on the left. The syntactic unit is shown on the right. The transient structure actually contains more units, and the units themselves contain more features, but everything has been shortened for illustrative purposes. The top row shows the lexical unit after the application of morphological and lexical constructions. The parsed string unambiguously allows for a decision to be made on the `lex-cat` value, and hence the value is set on the syntactic side. It is the functional construction that picks one of the `potential` `types` on the semantic side and fills its `value` attribute.

7.2.3 Discussion

Handling the many-to-many mapping problem in lexical class choice in principle also has other solutions. In particular, one could rely on the search process of FCG in order to branch into all possible lexical classes for a particular lexical item, and cut branches until only the one relevant for the current purpose, e.g. for the meaning produced, survives. This solution requires one lexical constructions for each possible lexical class and for each category. This leads to excessive branching in search. The solution presented in this chapter does not require branching in search and is thus more efficient. Another solution to this problem is to code the interaction of lexical items and lexical classes in a holistic fashion, which means that every possible combination of lexical class and lexical item is represented by exactly one construction. This solution does depend on branching of search, but demands grammar designers to hand code many constructions (for the case discussed in this chapter more than 30 combined constructions are needed). The solution presented in this chapter leads to a concise grammar, with much fewer constructions (less than 20 lexical and functional constructions).

7.3 Landmarks and complements – adverbial and prepositional constructions

Chapter 5 contains a number of examples of landmark and perspective marking in German spatial phrases. Most importantly I concluded that in spatial prepositional phrases the landmark is always part of the prepositional phrase. This contrasts with adverbs which allow landmarks to be expressed optionally using prepositional complements. Additionally, we have seen in Chapter 6 how spatial semantics chiefly relies on particular semantic operations and linking of semantic structure. Consequently, there are two important questions related to processing landmarks and perspective in adverbial and prepositional phrases: (1) how to deal with optional elements in a concise way, and (2) how to achieve linking. In this section I explore the solution to these problems for projective categories, in particular projective adverbs and projective prepositions.

Let us first consider the extension of projective prepositions using a landmark. The construction handling projective and absolute prepositions is called `angular-pp-phrase`. It has two constituents (see Table 7.3). The first constituent is one that has the semantic class `angular-relationship` and the syntactic class `angular-preposition` (see Table 7.2 and in particular the semantic and syntactic function attributes of all projective prepositional constructions `frontal-preposition`, `lateral-`

Table 7.3: Syntactic and semantic mappings of constructions governing prepositional and complemented adverbials. Syntactic and semantic functions and classes for the two constituents (c1 and c2) are shown. To the right the newly introduced syntactic and semantic function and classes are shown.

reverse columns and rows?

cxn	c1 syn-function	c1 syn-class	c2 syn-fn	c2 syn-class	syn-functions	syn-classes	examples
angular-pp-phrase		angular-preposition	referring-expression		adverbial		[vor] [der Kiste], [links] [der Kiste]
lateral-region-landmark-marked	adverbial	adverb	referring-expression		adverbial		[links] [von mir]
relative-region-perspective-marked	adverbial		referring-expression				[vor der kiste] [von mir aus], [links] [von mir aus], [links von der Kiste] [von mir aus]
cxn name	**c1 sem-function**	**c1 sem-class**	**c2 sem-fn**	**c2 sem-class**	**sem-functions**	**sem-classes**	**examples**
angular-pp-phrase		angular-relationship	reference		modifier	region, relative-region	[vor][der Kiste], [links][der Kiste]
lateral-region-landmark-marked	modifier	lateral-region	reference		modifier	region, relative-region	[links] [von mir]
relative-region-perspective-marked	modifier	relative-region	reference		modifier	region	[vor der Kiste] [von mir aus],[links] [von mir aus], [links von der Kiste] [von mir aus]

preposition and vertical-preposition) functional construction). The second constituent is the landmark. For the landmark there are only functional constraints. Whatever is supposed to act as the landmark needs to be some kind of reference (semantic) and referring-expression (syntactic). How is the linking precisely achieved? When we look at the prepositional construction in (7), we see that it features the variable ?landmark connected to the landmark slot of the respective operation. Since FCG cannot rely on variable names as they might change, the variable is repeated in the args feature, clearly marked in the attribute landmark. This means that the angular-pp-phrase construction can unify with this specific argument, which is for this purpose. Let us look at the template to understand the linking.

(12)
```
(def-phrasal-cxn angular-pp-phrase
 :constituents
 (def-constituent
  :syn-class angular-preposition
  :sem-class angular-relationship
  :args ((landmark ?landmark)))
 (def-constituent
  :syn-function referring-expression
  :sem-function reference
  :args ((ref ?landmark)))
  :syn-function adverbial
  :sem-function modifier)
```

Linking is achieved by explicitly unifying the corresponding args in the structure, using the variable ?landmark. The variable occurs both in the ?ref argument of the reference constituent and the landmark argument of the angular-relationship constituents.

For adverbs the linking with landmark works similar to prepositions. However, there are important differences in syntax between prepositions and adverbs. Let us consider the example of lateral adverbs. Lateral adverbs can be extended by landmarks, but the has to be marked using a *von* prepositional phrase. The construction handling the landmark augmentation of lateral adverbs is shown below. In addition to linking, this construction introduces the preposition *von*.

(13)
```
(def-phrasal-cxn lateral-adverb-landmark-marked
 :constituents
 (def-constituent
  :syn-function adverbial
  :syn-class lateral-adverb
  :sem-function modifier
```

```
 :sem-class lateral-region
 :args ((landmark ?landmark)))
(def-constituent
 :syn-function referring-expression
 :sem-function reference
 :args ((ref ?landmark))
 :preposition ”von”)
:syn-function adverbial
:sem-function modifier)
```

These two constructions show how to solve parts of the complexity puzzle of the interaction of syntax and semantics, i.e. the linking issue, while at the same time they deal with the syntactic differences for adverbs (in this case only lateral adverbs are discussed, but similar constructions exist for vertical and frontal adverbs) and prepositions. Moreover, the two constructions also prevent frontal adverbs from being landmark-augmented by any of the two constructions, since frontal adverbs do not have the `angular-preposition` potential. Also they cannot be extended using the lateral landmark marking scheme, since they are not of class `lateral-region`. Other constructions deal with *in* prepositional phrases and frontal adverbs.

7.4 Linking everything together – high-level phrasal constructions

The previous section looked at prepositional and adverbial constructions. But there is more. Particularly, there are important constructions which only care about the syntactic and semantic functions of their constituents, and hence are widely applicable and underlie the ability of the grammar to build and parse complex recursive utterances involving many complemented phrases. Phrasal constructions have two constituents. The unification of constituents is based on the actual-potential design pattern. The constructions require their constituents to provide certain semantic and syntactic function potentials, while providing new potentials for semantic and syntactic functions themselves. All of them also introduce a particular word order and a particular linking of the arguments of their constituents and the meaning they express. They internally link the arguments of constituents while providing new arguments themselves. Hence, in production these constructions express particular linkings in the semantic structure using a particular word order. Vice versa, in parsing they introduce links in the semantic structure when observing a particular word order of their functional constituents.

Table 7.4: Mapping of syntactic functions (phrasal constructions). All phrasal constructions have two constituents and all build hierarchical structure by subsuming the two constituents (*c1* and *c2*) into a new unit. Columns *c1 syn-fns* and *c2 syn-fns* show the syntactic function potential expected from constituents. The column *syn-fns* details the syntactic function potential of the new unit. All constructions shown here introduce word order and require the first constituent *c1* to meet the second constituent *c2*, i.e. *c1* has to be exactly before *c2*.

cxn	c1 syn-fn	c2 syn-fn	syn-fns	examples
adjectival–nominal–phrase	adjectival	nominal	nominal	[linke] [block]
determiner–nominal–phrase	determiner	nominal	referring-expression	[der] [block]
referring-expression-adverbial-phrase	referring-expression	adverbial	referring-expression	[der block] [links], [der block] [vor/an...]
preposition–referring-expression–phrase	preposition	referring-expression	adverbial	[an][der kiste]
possessive	referring-expression	referring-expression	referring-expression	[die linke Seite] [der Kiste]

The simplest example of such a construction is the `adjectival-nominal-phrase`, which allows agents to build large adjective noun phrases (see Steels 2011a for the original idea). Tables 7.4 and 7.5 detail the semantic and syntactic functions of the constituents of all phrasal constructions, as well as the syntactic and semantic function potentials they introduce. The adjectival nominal construction maps a constituent with syntactic function `adjectival` and semantic function `modifier` and a constituent with syntactic function nominal and semantic function identifier onto a new unit.Hence, it builds hierarchy by introducing a new unit with two subunits – namely its two constituents. This new unit has the semantic function potential of identifier and the syntactic function potential nominal. There are a number of functional constructions providing such semantic and syntactic functions. Both color and spatial adjectives provide the semantic function `modifier` and the syntactic function `adjectival`. The semantic function `identifier` and the syntactic function `nominal` are provided by nouns. Hence, when encountering such constituents in production, for instance because noun and adjective functional constructions have provided suitable constituents, the construction will form a new unit with semantic function nominal and introduce the German word order, where adjectives always come before the noun. This new structure can itself be considered functionally equal to nouns, as it features the same syntactic and semantic functions. It therefore can be subject to modification through other adjectives. Finally, units that have the semantic function `identifier` and the syntactic function `nominal`, can be extended by determiners through application of the `determiner-nominal-phrase`, which results in a unit that encapsulates the semantic function `reference` and the syntactic function `referring-expression` and provides for all examples, where nouns or adjective modified nouns are determined using an article.

???

This explains how adjectival noun phrases can be build, but how do adverbial complement phrases discussed in the previous section get linked to referring expressions? This is solved by `referring-expression-adverbial-phrase`, which links constituents with the semantic/syntactic function `reference/referring-expression` (example *der Block* or *der grüne Block* and `modifier/adverbial` (example *links*, *links von der Kiste*, *vor der Kiste …*) into a unit that not only syntactically introduces the word order, that the adverbial is behind the referring expression, but also links the meaning and adds the operation `apply-region-filter`. This construction, besides linking meaning of constituents, adds an operation that is applied to the output of the meaning of the adverbial phrase. Here, this operation is `apply-region-filter` which filters the context given a particular region. It is important to understand, that this particular construction can be so general only because all complements

Table 7.5: Mapping of semantic functions of phrasal constructions. For every construction the semantic function potential that needs to be present for the two constituents *c1* and *c2* is shown, as well as the new semantic function potential provided (*sem-fns*). Some of the constructions add additional meaning with more complicated argument linking properties. All others however link the `ref` argument of constituent two (*c2*) to the `source` argument of constituent one (*c1*).

cxn	c1 sem-fn	c2 sem-fn	sem-fns	operation	examples
adjectival–nominal–phrase	modifier	identifier	identifier		[linke] [block]
determiner–nominal–phrase	determiner	identifier	reference		[der] [block]
preposition–referring-expression–phrase	relation	reference	modifier		[an] [der kiste]
referring-expression-adverbial-phrase	reference	modifier	reference	apply-region-filter	[der block] [links], [der block] [vor/an...]
possessive	reference	reference	ref	possessive	[die linke Seite] [der Kiste]

in the grammar compute regions. In other words all adverbial complements always denote a spatial region, be they prepositional phrases or adverbials or complemented adverbials. The `referring-expression-adverbial-phrase` construction, that handles all adverbial complements of determined nominal phrases, only needs to care about modifiers and hence, its generality is based on the fact that semantically all adverbial complements in this grammar compute regions while observing a particular word order. If there would be other possible complements, this construction would also need to specialize on the semantic type of the constituents.

7.5 Handling case

Case and gender agreement in German is an example of a highly distributed information processing task. The constraints on these syntactic features are contributed by many different constructions and thus have to be incrementally integrated in order to produce grammatical utterances in German. For instance, the grammatical gender of a prepositional determined adjective noun phrase is determined by the noun, as shown in the following example (*Block*, masculine).

(14) hinter dem linken Block
behind.PREP the.DAT left.DAT block.DAT
'behind the left block'

Alternatively, the case is governed by the preposition (*hinter*, requires dative). The determiner (*der*) and the adjective (*link*) have to be case and gender marked according to the information provided from these different sources, namely, both the determiner and adjective are used in their masculine dative forms (*dem* and *linken*). In other words, the concrete form of a projective adjective can only be fixed after the complete syntactic structure is processed. Along the way, information about which case to use (coming from the preposition) and about the gender (from the noun) need to be integrated. Consequently, the grammar needs to be set up such that sets of highly dependent constructions can interact for allowing a distributed decision on which forms to use when expressing a particular meaning. This includes mechanisms for (1) representing the state of information including its uncertainty, (2) moving information around in order to facilitate decisions and spread their effect, and (3) ways to postpone decisions until enough information is accumulated. The solutions presented for these problems naturally mirror the techniques discussed in the previous section. Logic variables embedded in feature matrices are used to represent uncertainty, percolation for moving information around and constructions of a particular type in order to postpone decisions.

7.5.1 Representing the state of information

Distinctive feature matrices (see van Trijp 2011) are a means to represent the current, possibly indecisive state of information in processing. They allow different constructions to independently contribute constraints on values of the syntactic, case and gender features until enough information has been collected. Hence, feature matrices function similarly to the logic variable used for representing uncertainty in the previous section, as they are a technique for accumulating information contributed by different constructions. Distinctive feature matrices extend the concept of logic variables and allow for the representation of dependencies between features in processing.

The way lexical items interact with the case gender agreement system is determined in part by the lexical item and in part by the word class. Nouns, for instance, have a particular gender and always need to be marked for case, which is governed by prepositions. Adjectives and articles agree in case and number with the phrase in which they are embedded, specifically with the noun. Consequently, the state of information for some word classes is initially constrained. While adjectives and articles have no constraints on case and gender, nouns already provide information about their gender, and prepositions about the required case. Distinctive feature matrices allow for the representation of such different states of information in the transient structure in a unified way by explicitly representing all combinations of possible feature values in a matrix. For our German example, this information is captured in a two dimensional matrix, where columns reflect the four German cases, and rows reflect the three grammatical genders. Every field in the matrix corresponds to a particular combination of case and gender, such as accusative-masculine, and every field can either be explicitly excluded (i.e. marked with a '–'), selected (i.e. marked by a '+') or in an unknown state of information, which is represented using variables i.e. marked with a '?').

Figure 7.6 shows the state of the transient structure after the application of lexical and functional constructions. It can be seen how the different states of information for articles, adjectives, prepositions and nouns are technically represented. The feature matrices for the spatial adjective (`spatial-adjective-unit-334`) and for the article (`article-unit-334`) are completely filled with variables. On the other hand, the feature matrix for the frontal preposition (`frontal-preposition-unit-93`) features a '–' everywhere but in the column representing the dative case, namely, the case it requires. On the other hand, the noun (`noun-unit-334`) is categorized based on its gender, and the feature matrix consequently has variables in the row for masculine and excludes all other fields.

7.5.2 Percolation and agreement – moving information around and unification

Given the setup of initial information by lexical and functional constructions, all subsequently applied constructions have to be able to move information around and to further constrain the information. Movement of information is done using percolation, and unification of feature matrices for agreement automatically

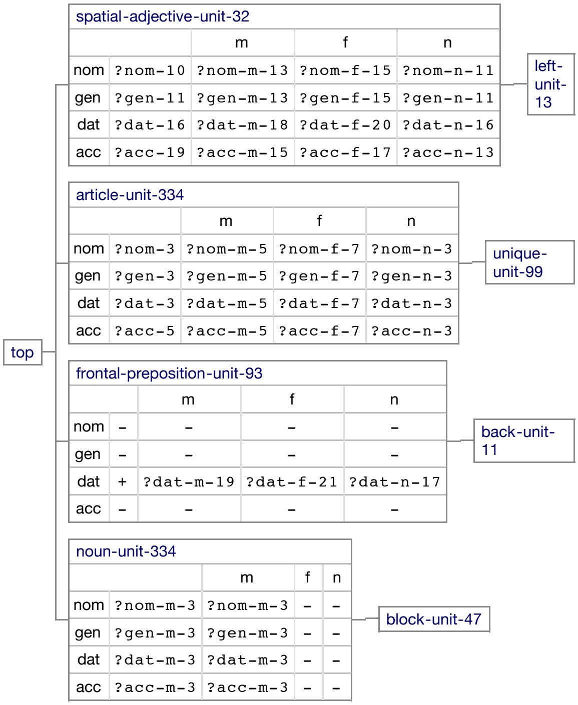

Figure 7.6: Transient structure after the application of lexical and functional constructions for production of *hinter dem linken Block* ('behind the left block'). For simplification, each unit is only shown with its distinctive feature matrix for case/gender agreement, if present. Furthermore, the feature matrices of the lexical units are identical to those of their parent units and are thus also not shown.

constrains the values in the feature matrices further and further.

Both percolation and unification are used together, for instance, by the `adjectival-nominal` construction (see Figure 7.7). In our example, this construction handles the adjective (`spatial-adjective-unit-334`) and the noun (`noun-unit-334`) as constituents. Apart from introducing German word order, this construction unifies the feature matrix of the adjective and the noun, which automatically constrains the gender possibilities for the adjective, in this case to masculine. In fact, through unification the two feature matrices are the same after the application of the `adjectival-nominal` constructions. Moreover, the newly created parent unit (`adjectival-nominal-ph` percolates this matrix up. This process is subsequently repeated, this time by the `determiner-nominal` construction, which has the same effect but this time with its

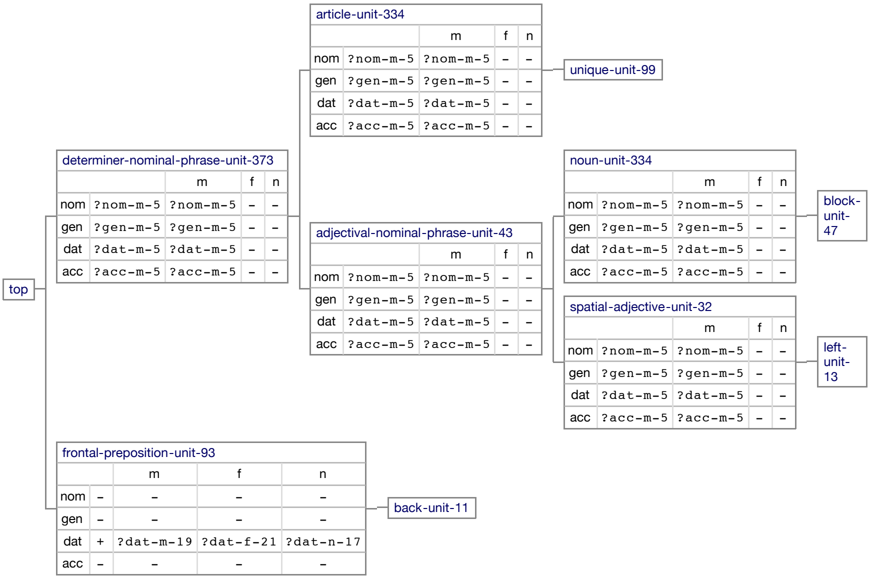

Figure 7.7: Gender agreement between the article, adjective and noun are enforced by the adjectival-nominal and determiner-nominal-phrase constructions applied to the transient structure in Figure 7.6.

constituents being the article and the adjectival-nominal phrase, which also constrains the article to be masculine. Percolation and unification have essentially established the agreement between the article, the adjective and the noun, while at the same time spreading the information about gender.

After the application of these two constructions, the decision on case is still missing. Case is provided by the angular preposition, and agreement between the preposition and the determined-nominal-phrase is established by the `angular-pp-phrase` (see Figure 7.8). The `angular-pp-phrase` technically behaves very similarly to the the `determiner-nominal` and the `adjectival-nominal` constructions: it unifies the feature matrices of its two constituents (`frontal-preposition-unit-93` and `determiner-nominal-phrase-unit-373`) However, the effect is quite different in that now the feature matrix of the article, the adjective and the noun is further constrained in terms of case. Consequently, case and gender of this particular phrase are ultimately decided.

For some phrases case is not established by prepositions but rather by the general grammatical structure. A prime example is when the utterance itself only consists of a noun phrase, which then needs to be marked by the nominative case. For example when meanings for utterances such as *der linke Block* (‘the block to

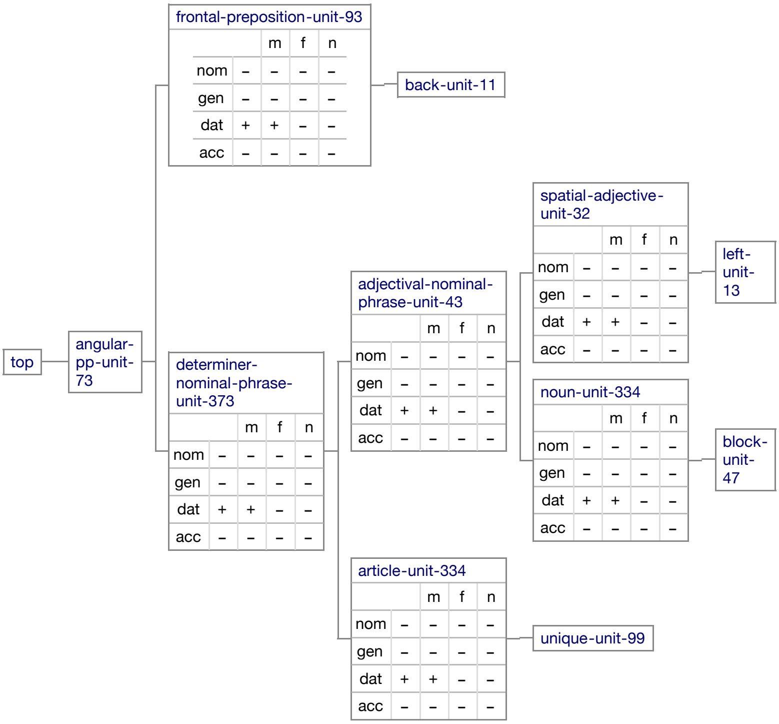

Figure 7.8: Case agreement after applying the angular-pp-phrase construction to the transient structure from Figure 7.7 while producing "hinter dem linken block".

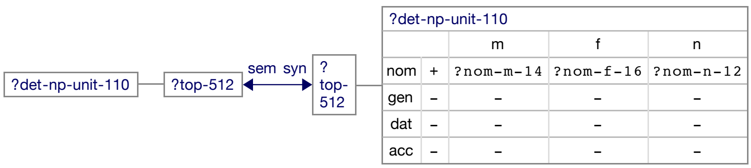

Figure 7.9: The referring-expression construction sets the case of a single determined-noun-phrase unit to nominative.

the left') are expressed, then there is no preposition that determines the case of the whole phrase by agreement. Rather, the `referring-expression` construction (see Figure 7.9) introduces the nominative case by unifying the feature matrix of the determined-noun-phrase unit with a matrix constraining the case to nominative.

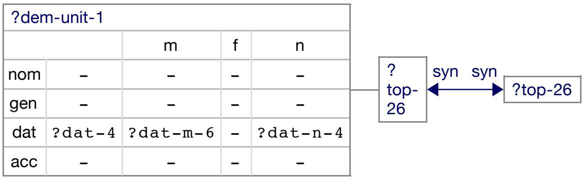

Figure 7.10: Distinctive feature matrix of the morphological construction that maps the string *dem* to masculine or neuter and dative articles. Note that since this is a morphological construction, both poles of the construction apply to the syntactic pole of a transient structure.

7.5.3 Postponing decisions

After the application of the `angular-pp-phrase` construction, all necessary information has been accumulated. Case and gender are decided, and hence all syntactic features for the particular word class in question are available to allow subsequent constructions to be able to decide the word form to be used. Morphological constructions are used here to represent this relationship between syntactic features and word forms. For example, for determiners there are six different articles in German that unevenly cover the 12 possible case-gender combinations, as shown in the chart below:

	m	f	n
nom	*der*	*die*	*das*
gen	*des*	*der*	*des*
dat	***dem***	*der*	***dem***
acc	*den*	*die*	*das*

For each of these forms, a separate morphological construction exists which decides on the form used to express the article based on the lexical class and the case-gender feature matrix. An example of such a morphological construction is shown in Figure 7.10. Since this construction has a variable in the dative masculine field, it matches with unit `unique-unit-99` in Figure 7.8. Similarly, other morphological entries add the strings *linken* to the `block-unit-47`, *Block* to the `block-unit-47` and *hinter* to `back-unit-11`.

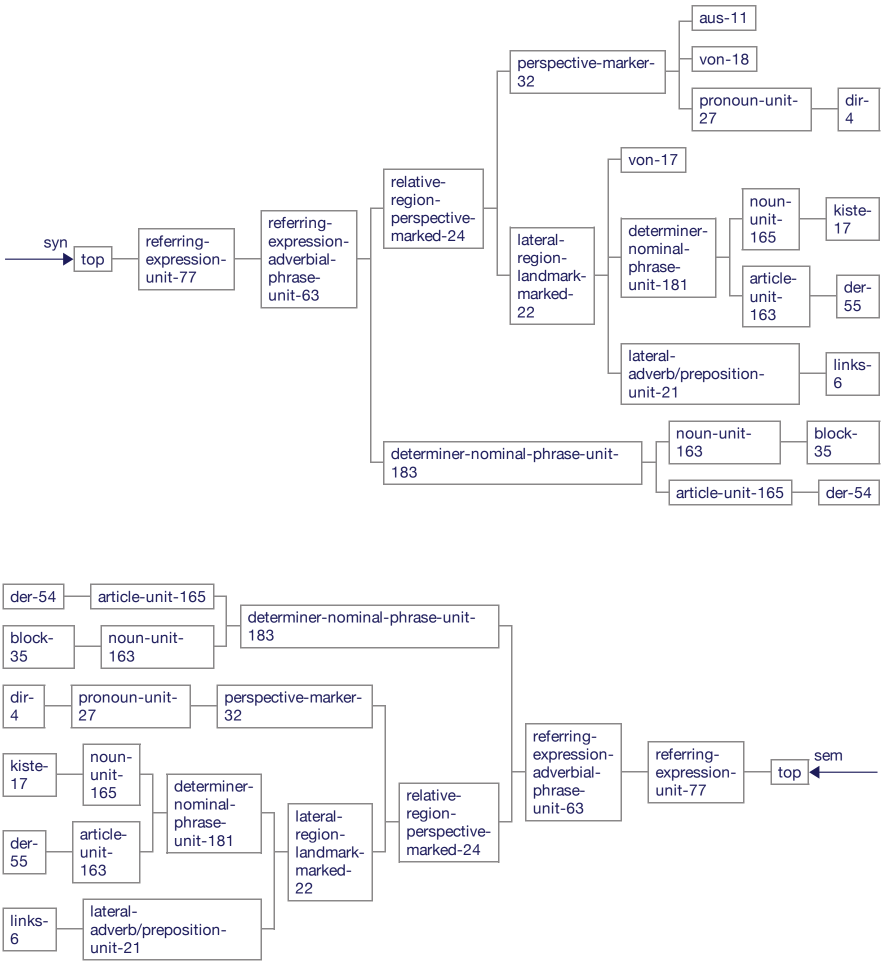

Figure 7.11: Parts of a final transient structure. The top shows the syntactic pole. The bottom shows the semantic pole.

7.6 Summary

Problems of intertwined information processing across multiple constructions and across different parts of transient structures often appear when dealing with complex, real world language. This chapter detailed how to tackle such problems using (1) adequate information representation techniques, such as logic variables, feature matrices and disjunctive potentials, (2) percolation for distributing information in the transient structure, and (3) special constructions which are needed to help postpone decisions until the state of information is ready. The techniques have proven to be sufficient for handling problems of syntactic indeterminacy, e.g. morphology and lexical class choice in German locative phrases. The discussed design patterns allow grammar designers to spread information processing across many constructions, leading to concise grammars, while facilitating efficient processing.

The grammar is powerful enough to deal with German locative phrases that include spatial relations, landmarks and even perspective. Figure 7.5.3 shows the final transient structure when an agent parses the phrase *der Block links von der Kiste von dir aus* ('the block to the left of the box from your perspective').

Part III

Conclusion

Bibliography

Arbib, M. A. 2002. The Mirror System, Imitation, and the Evolution of Language. In K. Dautenhahn & C. Nehaniv (eds.), *Imitation in Animals and Artifacts*, 229–280. The MIT Press.

Bailey, D., J. Feldman & S. Narayanan. 1997. Modeling embodied lexical development. In *Proceedings of the nineteenth annual conference of the cognitive science society: August 7-10, 1997, stanford university*, 19. Lawrence Erlbaum Associates.

Bateman, J. A. 2010. Situating spatial language and the role of ontology: Issues and outlook. *Language and Linguistics Compass* 4(8). 639–664. doi:10.1111/j.1749-818X.2010.00226.x. http://dx.doi.org/10.1111/j.1749-818X.2010.00226.x.

Bateman, J. A., J. Hois, R. Ross & T. Tenbrink. 2010. A linguistic ontology of space for natural language processing. *Artificial Intelligence* 174(14). 1027 – 1071. doi:DOI:10.1016/j.artint.2010.05.008. http://www.sciencedirect.com/science/article/B6TYF-506RCV8-1/2/9e417a4d8dae6e26102301f49e205489.

Bateman, J. A., T. Tenbrink & S. Farrar. 2007. The Role of Conceptual and Linguistic Ontologies in Interpreting Spatial Discourse. *Discourse Processes* 44(3). 175–212.

Belpaeme, T. & J. Bleys. 2007. Language, Perceptual Categories and their Interaction: Insights from Computational Modelling. In C. Lyon, C. Nehaniv & A. Cangelosi (eds.), *Emergence of communication and language*, 339–353. London, UK: Springer.

Bickerton, D. 1999. *Language and species.* Chicago University Press.

Billard, A. & K. Dautenhahn. 1998. Grounding communication in autonomous robots: An experimental study. *Robotics and Autonomous Systems* 24(1). 71–79.

Bleys, J. 2010. *Language strategies for the domain of colour.* Brussels, Belgium: Vrije Universiteit Brussels (VUB) dissertation.

Bleys, J., M. Loetzsch, M. Spranger & L. Steels. 2009. The Grounded Color Naming Game. In *Proceedings of the 18th ieee international symposium on robot and human interactive communication (ro-man 2009)*, .

Van den Broeck, W. 2008. Constraint-based compositional semantics. In A. D. M. Smith, K. Smith & R. Ferrer i Chancho (eds.), *The evolution of language (evolang 7)*, Singapore: World Scientific.

Camazine, S., J. L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraula & E. Bonabeau. 2003. *Self-organization in biological systems*. Princeton University Press.

Carlson-Radvansky, L. A. & G. A. Radvansky. 1996. The influence of functional relations on spatial term selection. *Psychological Science* 7(1). 56–60.

Carpenter, M., K. Nagell, M. Tomasello, G. Butterworth & C. Moore. 1998. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child development* 63(4).

Christiansen, M. H., R. Dale, M. R. Ellefson & C. M. Conway. 2001. The role of sequential learning in language evolution: Computational and experimental studies. In A. Cangelosi & D. Parisi (eds.), *Simulating the Evolution of Language*, 165–188. Springer.

Cohn, A. G. & S. M. Hazarika. 2001. Qualitative spatial representation and reasoning: An overview. *Fundamenta Informaticae* 46(1-2). 1–29. http://portal.acm.org/citation.cfm?id=1219982.1219984.

Coradeschi, S. & A. Saffiotti. 2003. An introduction to the anchoring problem. *Robotics and Autonomous Systems* 43(2-3). 85–96.

Coventry, K. & S. Garrod (eds.). 2004. *Saying, Seeing and Acting. The Psychological Semantics of Spatial Prepositions* (Essays in Cognitive Psychology). Lawrence Erlbaum Associates.

Coventry, K., M. Prat-Sala & L. Richards. 2001. The interplay between geometry and function in the comprehension of over, under, above, and below. *Journal of Memory and Language* 44(3). 376 – 398. doi:DOI:10.1006/jmla.2000.2742. http://www.sciencedirect.com/science/article/B6WK4-457VF3J-1C/2/aeb0c9d7b26f70d42c12745494816549.

Coventry, K. R., A. Cangelosi, R. Rajapakse, A. Bacon, S. Newstead, D. Joyce & L. Richards. 2005. Spatial Prepositions and Vague Quantifiers: Implementing the Functional Geometric Framework. In C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel & T. Barkowsky (eds.), *Spatial cognition iv. reasoning, action, interaction*, vol. 3343 (Lecture Notes in Computer Science), 98–110. Springer.

Croft, W. 2001. *Radical construction grammar*. Oxford University Press.

De Beule, J. & L. Steels. 2005. Hierarchy in fluid construction grammars. *KI 2005: Advances in Artificial Intelligence* 1–15.

Dunbar, R. 1998. Theory of mind and the evolution of language. In *Approaches to the Evolution of Language: Social and Cognitive Bases*, 92–110. Cambridge University Press.

Dunbar, R. 2003. The social brain: Mind, language, and society in evolutionary perspective. *Annual Review of Anthropology* 32. 163–181.

Eschenbach, C. 1999. Geometric structures of frames of reference and natural language semantics. *Spatial Cognition and Computation* 1(4). 329–348.

Eschenbach, C. 2004. Contextual, Functional, and Geometric Components in the Semantics of Projective Terms. In L. Carlson & E. van der Zee (eds.), *Functional Features in Language and Space*, vol. 2 (Explorations in Language and Space), 71–93. Oxford University Press.

Evans, N. & S. C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(05). 429–448. doi:10.1017/S0140525X0999094X.

Freksa, C. 1991. Qualitative Spatial Reasoning. In D. M. Mark & A. U. Frank (eds.), *Cognitive and linguistic aspects of geographic space*, 361–372. Springer.

Gapp, K. 1994. Basic meanings of spatial relations: Computation and evaluation in 3d space. In *Proceedings of aaai-94*, 1393–1398.

Gapp, K. 1995. Angle, distance, shape, and their relationship to projective relations. In *Proceedings of the 17th annual conference of the cognitive science society*, 112–117.

Garrod, S. & G. Doherty. 1994. Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition* 53(3). 181–215.

Gerasymova, K. & M. Spranger. 2010. Acquisition of grammar in autonomous artificial systems. In H. Coelho, R. Studer & M. Woolridge (eds.), *Proceedings of the 19th european conference on artificial intelligence (ecai-2010)*, 923–928. IOS Press.

Gerasymova, K. & M. Spranger. 2012. An experiment in temporal language learning. In L. Steels & M. Hild (eds.), *Language grounding in robots*, 237–254. Springer.

Goldberg, A. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago University Press.

Grabowski, J. & P. Weiss. 1996. The prepositional inventory of languages: A factor that affects comprehension of spatial prepositions. *Language Sciences* 18(1-2). 19 – 35. doi:DOI:10.1016/0388-0001(96)00005-8. http://www.sciencedirect.com/science/article/B6VD2-3VW1RDD-1/2/c20d6b8cabef41affbf397ee2ff84d55. Contrastive Semantics and Pragmatics, Volume I: Meanings and Representations.

Haddock, N. J. 1989. Computational models of incremental semantic interpretation. *Language and Cognitive Processes* 4(3). 337–368.

Hall, M. M. & C. B. Jones. 2008. Quantifying spatial prepositions: an experimental study. In *Proceedings of the 16th acm sigspatial international conference on*

advances in geographic information systems, 1–4. ACM.
Herskovits, A. 1986. *Language and spatial cognition* (Studies in Natural Language Processing). Cambridge University Press.
Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization* (Cambridge Textbooks in Linguistics). Cambridge University Press.
Jackendoff, R. 1999. Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences* 3(7). 272–279.
Johnson-Laird, P. N. 1977. Procedural semantics. *Cognition* 5(3). 189–214.
Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering* 82(1). 35–45.
Kauffman, S. A. 1993. *The Origins of Order: Self-organization and Selection in Evolution*, vol. 209. Oxford University Press.
Kelleher, J. D. & F. J. Costello. 2009. Applying computational models of spatial prepositions to visually situated dialog. *Computational Linguistics* 35(2). 271–306. doi:10.1162/coli.06-78-prep14. http://www.mitpressjournals.org/doi/abs/10.1162/coli.06-78-prep14.
Klabunde, R. 1999. Logic—Based Choice of Projective Terms. *KI-99: Advances in Artificial Intelligence* 697–697.
Levinson, S. C. 1996. Language and space. *Annual review of Anthropology* 25(1). 353–382.
Levinson, S. C. 2003. *Space in Language and Cognition: Explorations in Cognitive Diversity* (Language, Culture and Cognition 5). Cambridge University Press.
Levinson, S. C. & D. Wilkins. 2006. *Grammars of Space*. Cambridge University Press.
Lloyd, S.P. 1982. Least squares quantization in PCM. *IEEE Transactions on Information Theory* 28(2). 129–137.
Maddieson, I. 1984. *Patterns of sounds*. Cambridge University Press.
Miller, G. A. & P. N. Johnson-Laird. 1976. *Language and perception*. Belknap Press.
Moratz, R. & T. Tenbrink. 2006. Spatial reference in linguistic human-robot interaction: Iterative, empirically supported development of a model of projective relations. *Spatial Cognition & Computation* 6(1). 63–107.
Narayanan, S. 1999. Moving right along: a computational model of metaphoric reasoning about events. In *Proceedings of the national conference on artificial intelligence (aaai'99)*, 121–127. American Association for Artificial Intelligence Menlo Park, CA, USA.
Oudeyer, P. Y. 2005. The self-organization of speech sounds. *Journal of Theoretical Biology* 233(3). 435–449.

Pickering, M. J. & S. Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27(2). 169–190.

Pope, M. K. 1952. *From latin to modern french with especial consideration of anglo-norman. phonology and morphology*. Manchester University Press.

Regier, T. & L. A. Carlson. 2001. Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General* 130(2). 273–298.

Rizzolatti, G. & M. A. Arbib. 1998. Language within our grasp. *Trends in neurosciences* 21(5). 188–194.

Rosch, E. 1978. Principles of categorization. In E. Rosch & B.B. Lloyd (eds.), *Cognition and categorization*, 27–48. Hillsdale, NJ: Lawrence Erlbaum.

Roy, D. 2005. Semiotic schemas: A framework for grounding language in action and perception. *Artificial Intelligence* 167(1-2). 170–205.

Schoenemann, P. 1999. Syntax as an emergent characteristic of the evolution of semantic complexity. *Minds and Machines* 9(3). 309–346.

Siskind, J. M. 2001. Grounding the Lexical Semantics of Verbs in Visual Perception using Force Dynamics and Event Logic. *Journal of Artificial Intelligence Research* 15. 31–90.

Smith, K., S. Kirby & H. Brighton. 2003. Iterated learning: A framework for the emergence of language. *Artificial Life* 9(4). 371–386.

Sperber, D. & D. Wilson. 1986. *Relevance: Communication and cognition*. Harvard University Press.

Spranger, M. 2008. *World models for grounded language games*. Humboldt-Universität zu Berlin German diplom thesis.

Spranger, M. & M. Loetzsch. 2011. Syntactic Indeterminacy and Semantic Ambiguity: A Case Study for German Spatial Phrases. In L. Steels (ed.), *Design patterns in Fluid Construction Grammar*, vol. 11 (Constructional Approaches to Language), 265–298. John Benjamins.

Spranger, M., M. Loetzsch & S. Pauw. 2010. Open-ended Grounded Semantics. In H. Coelho, R. Studer & M. Woolridge (eds.), *Proceedings of the 19th european conference on artificial intelligence (ecai 2010)*, vol. Frontiers in Artificial Intelligence and Applications 215, 929–934. IOS Press.

Spranger, M., M. Loetzsch & L. Steels. 2012a. A Perceptual System for Language Game Experiments. In L. Steels & M. Hild (eds.), *Language Grounding in Robots*, 89–110. Springer.

Spranger, M., S. Pauw, M. Loetzsch & L. Steels. 2012b. Open-ended Procedural Semantics. In L. Steels & M. Hild (eds.), *Language Grounding in Robots*, 153–172. Springer.

Steels, L. 1997. Constructing and sharing perceptual distinctions. In M. van Someren & G. Widmer (eds.), *Proceedings of the ninth european conference on machine learning*, 4–13. Springer.

Steels, L. 1998. Synthesising the Origins of Language and Meaning Using Co-evolution, Self-organisation and Level formation. In J. Hurford, C. Knight & M. Studdert-Kennedy (eds.), *Approaches to the evolution of language: Social and cognitive bases*, 384–404. Edinburgh University Press.

Steels, L. 2000a. The emergence of grammar in communicating autonomous robotic agents. In W. Horn (ed.), *Ecai 2000: Proceedings of the 14th european conference on artificial intelligence*, 764–769. IOS Publishing.

Steels, L. 2000b. Language as a complex adaptive system. In M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J.J. Merelo & H-P. Schwefel (eds.), *Proceedings of the 6th international conference on parallel problem solving from nature*, vol. 1917 (Lecture Notes in Computer Science), 17–26. Springer.

Steels, L. 2001. Language games for autonomous robots. *IEEE Intelligent systems* 16–22.

Steels, L. 2005. The emergence and evolution of linguistic structure: From lexical to grammatical communication systems. *Connection Science* 17(3-4). 213–230.

Steels, L. 2007. The Recruitment Theory of Language Origins. In C. Lyon, C. Nehaniv & A. Cangelosi (eds.), *The Emergence of Communication and Language*, 129–151. Springer.

Steels, L. 2008. The symbol grounding problem has been solved. so what's next? In M. de Vega (ed.), *Symbols and embodiment: Debates on meaning and cognition*, Oxford University Press.

Steels, L. 2009. Cognition and social dynamics play a major role in the formation of grammar. In D. Bickerton & E. Szathmáry (eds.), *Biological foundations and origin of syntax*, vol. 3 (Strungmann Forum Reports), 345–368. MIT Press.

Steels, L. 2011a. A Design Pattern For Phrasal Constructions. In L. Steels (ed.), *Design Patterns in Fluid Construction Grammar*, John Benjamins.

Steels, L. (ed.). 2011b. *Design patterns in Fluid Construction Grammar.* John Benjamins.

Steels, L. 2011c. Self-organization and Selection in Cultural Language Evolution. In L. Steels (ed.), *Experiments in Cultural Language Evolution*, John Benjamins.

Steels, L. (ed.). 2012. *Computational Issues in Fluid Construction Grammar* (Lecture Notes in AI). Springer Verlag.

Steels, L. & J.-C. Baillie. 2003. Shared grounding of event descriptions by autonomous robots. *Robotics and Autonomous Systems* 43(2-3). 163–173.

Steels, L. & T. Belpaeme. 2005. Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences* 28. 469–529.

Steels, L. & J. Bleys. 2005. Planning what to say: Second order semantics for fluid construction grammars. In A. Bugarin Diz & J. Santos Reyes (eds.), *Proceedings of caepia 05.* (Lecture Notes in AI), Springer.

Steels, L. & J. De Beule. 2006. Unify and merge in Fluid Construction Grammar. In P. Vogt, Y. Sugita, E. Tuci & C. Nehaniv (eds.), *Symbol grounding and beyond: Proceedings of the third international workshop on the emergence and evolution of linguistic communication* (LNAI 4211), 197–223. Springer.

Steels, L., J. De Beule & N. Neubauer. 2005. Linking in fluid construction grammars. In *Transactions of the belgian royal society of arts and sciences*, 11–18.

Steels, L. & F. Kaplan. 1998. Stochasticity as a source of innovation in language games. In Christoph Adami, Richard K. Belew, Hiroaki Kitano & Charles E. Taylor (eds.), *Proceedings of the sixth international conference on artificial life*, The MIT Press.

Steels, L. & F. Kaplan. 2002. Bootstrapping grounded word semantics. In T. Briscoe (ed.), *Linguistic evolution through language acquisition: formal and computational models*, 53–74. Cambridge University Press.

Steels, L. & M. Spranger. 2008. The robot in the mirror. *Connection Science* 20(4). 337–358.

Steels, L. & M. Spranger. 2012. Emergent mirror systems for body language. In L. Steels (ed.), *Experiments in Cultural Language Evolution*, 87—109. John Benjamins.

Steels, L. & P. Wellens. 2006. How grammar emerges to dampen combinatorial search in parsing. In P. Vogt, Y. Sugita, E. Tuci & C. Nehaniv (eds.), *Symbol grounding and beyond: Proceedings of the third international workshop on the emergence and evolution of linguistic communication* (LNAI 4211), 76–88. Springer.

Talmy, L. 2000. *Toward a cognitive semantics. Vol. 2: Typology and process in concept structuring.* The MIT Press.

Tenbrink, T. 2005a. Identifying objects on the basis of spatial contrast: an empirical study. *Spatial Cognition IV. Reasoning, Action, and Interaction* 124–146.

Tenbrink, T. 2005b. *Localising objects and events: Discoursal applicability conditions for spatiotemporal expressions in English and German*: University of Bremen dissertation.

Tenbrink, T. 2007. *Space, time, and the use of language: An investigation of relationships*, vol. 36 (Cognitive Linguistics Research). Berlin, DE: Walter de

Gruyter.

Tinbergen, N. 1963. On aims and methods of ethology. *Zeitschrift für Tierpsychologie* 20(4). 410–433.

Tomasello, M. 1992. The social bases of language acquisition. *Social Development* 1(1). 67–87.

Tomasello, M. 1995. Joint attention as social cognition. In Chris Moore & Philip J. Dunham (eds.), *Joint attention: Its origins and role in development*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Tomasello, M., M. Carpenter, J. Call, T. Behne & H. Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28(05). 675–691.

Traugott, Elizabeth Closs & Bernd Heine (eds.). 1991. *Approaches to grammaticalization*, vol. 1 & 2 (Typologogical Studies in Language 19). John Benjamins.

van Trijp, R. 2008. *Analogy and multi-level selection in the formation of a case grammar. a case study in fluid construction grammar.* Antwerp: University of Antwerp dissertation.

van Trijp, R. 2011. Feature matrices and agreement: A case study for German case. In L. Steels (ed.), *Design Patterns in Fluid Construction Grammar*, 205–235. John Benjamins.

Vandeloise, C. 1991. *Spatial prepositions: A case study from French.* University of Chicago Press.

Vogt, P. 2002. Bootstrapping grounded symbols by minimal autonomous robots. *Evolution of Communication* 4(1). 87–116.

Winograd, T. 1971. *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language.*: Massachusetts Institute of Technology dissertation.

Worden, R. 1998. The evolution of language from social intelligence. In J. Hurford, M. Studdert-Kennedy & C. Knight (eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases*, 148–166. Cambridge University Press.

Wunderlich, D. & M. Herweg. 1991. Lokale und Direktionale. In A. Stechow & D. Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, De Gruyter.