Year: 2009

Modelling the Role of Pragmatic Plasticity in the Evolution of Linguistic Communication

Hoefler, S H

Modelling the Role of Pragmatic Plasticity in the Evolution of Linguistic Communication

Stefan H. Hoefler
MA UZH

A thesis submitted in fulfilment of requirements for the degree of Doctor of Philosophy

to

Linguistics and English Language
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

February 2009

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Stefan H. Hoefler

Abstract

For a long time, human language has been assumed to be genetically determined and therefore the product of biological evolution. It is only within the last decade that researchers have begun to investigate more closely the domain-general cognitive mechanisms of cultural evolution as an alternative explanation for the origins of language. Most of this more recent work focuses on the role of imperfect cultural transmission and abstracts away from the mechanisms of communication. Specifically, models developed to study the cultural evolution of language—both theoretical and computational—often tacitly assume that linguistic signals fully specify the meaning they communicate. They imply that ignoring the fact that this is not the case in actual language use is a justified idealisation which can be made without significant consequences. In this thesis, I show that by making this idealisation, we miss out on the extensive explanatory potential of an empirically attested property of language: its pragmatic plasticity. The meaning that a signal comes to communicate in a specific context usually differs to a certain degree from its conventional meaning. This thesis (i) introduces a model of the cultural evolution of language that acknowledges and incorporates the fact that communication exhibits pragmatic plasticity and (ii) explores the explanatory potential of this fact with regard to language evolution.

The thesis falls into two parts. In the first part, I develop the model conceptually. I begin by analysing the components of extant models of general cultural evolution and discuss how models of language change and linguistic evolution map onto them. Innovative use is identified as the motor of cultural evolution. I then conceptualise the cognitive mechanisms underlying innovative language use and argue that they originate in pre-linguistic forms of ostensive-inferential communication. In a next step, the identified mechanisms are employed to provide a unified account of the two main explananda of evolutionary linguistics, the emergence of symbolism and the emergence of grammar. Finally, I discuss the implications of the presented analysis for the so-called proto-language debate. In the second part of the thesis, I propose a computational implementation of the developed conceptual model. This computational implementation allows for the simulation of the cultural emergence and evolution of symbolic communication and provides a laboratory-like environment to study individual aspects of this process. I employ such computer simulations to explore the role that pragmatic plasticity plays in the development of the expressivity, signal economy and ambiguity of emerging and evolving symbolic communication systems.

As its main contribution to the study of language evolution, this thesis shows that a model of linguistic cultural evolution that incorporates the notion of pragmatic plasticity has the potential to explain two crucial evolutionary puzzles, namely (i) how language can emerge from no language, and (ii) how language can come to exhibit the appearance of design for communication. The proposed usage-based model of language evolution bridges the evolutionary gap between no language and language by identifying ostensive-inferential communication as the continual aspect present in both stages, and by demonstrating that the cognitive mechanisms involved in ostensive-inferential communication are sufficient for the transition from one stage to the other.

Acknowledgements

I wish to thank my supervisors, Simon Kirby and Jim Hurford, for the academic support and good advice they gave me during the preparation and completion of this thesis. The Cogito Foundation deserves acknowledgement for their generous financial support: without it, this project could not have been realised. For computing advice, my thanks go to Cedric Macmartin, Mike Bennett, Morag Brown and Eddie Dubourg.

I am also grateful to my colleages at the Language Evolution and Computation research unit and at the department of Linguistics, who, over the years, have become so numerous that they cannot all be mentioned here by name; special thanks for their helpful comments on various bits of the thesis go to Dan Dediu, Dave Hawkey, Anna Kinsella, Andrew Smith and Monica Tamariz.

I would also like to thank Mirjam Oberholzer, for helping me to keep things in perspective, and James Clawson, for sharing his invaluable LaTeX expertise.

Finally, I want to thank my parents and my grandparents, who worked hard to make my education possible, who taught me the value of independent thinking, and who never stopped encouraging and supporting me.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

> *"Good Morning!" said Bilbo, and he meant it. The sun was shining, and the grass was very green. But Gandalf looked at him from under long bushy eyebrows that stuck out further than the brim of his shady hat. "What do you mean?" he said. "Do you wish me a good morning, or mean that it is a good morning whether I want it or not; or that you feel good this morning; or that it is a morning to be good on?" "All of them at once," said Bilbo.*
>
> —J. R. R. Tolkien: The Hobbit.

Human language is the most expressive, most adaptive, and most efficient communication system in the animal world. Its story is one of unprecedented success, one that so far has not been repeated in any other species. The uniqueness of human language and the pivotal role it plays in human civilisation—and in what defines us as human in the first place—must almost inevitably evoke a desire to investigate its origins.

The existence of human language gives rise to two evolutionary puzzles: a puzzle of emergence and a puzzle of design. The *emergence puzzle* concerns the simple fact that language must have emerged from no language, that is, from a stage where our ancestors did not possess any form of linguistic or even just symbolic communication yet. The study of language evolution investigates how this gap between no language on the one side and human language as we have it today on the other side can be bridged. The *design puzzle* refers to the fact that language is not just some random communication system but that it appears to exhibit actual design for its function in communication: many of its features—though not all of them—seem to be tailored specifically for the the transmission of propositional information via a serial channel (Pinker and Bloom 1990:713). How did language come to be like this? Ultimately, any account of the origins of language must explain (at least) these two evolutionary puzzles: the emergence puzzle and the design puzzle.

1.1 Evolutionary linguistics

The field commonly labelled as "evolutionary linguistics" in fact comprises two quite different research enterprises. The reason for this division, which often goes unnoticed and is the cause of much miscommunication, is the fact that linguists do not agree on what language is in the first place. The disagreement evolves around the so-called innateness hypothesis, introduced by Chomsky (1959, 1965), which assumes that language, or at least some core part of it, is a domain-specific, genetically encoded module of the mind. Proponents of the innateness hypothesis view language as a component of human biology. In contrast, linguists who do not subscribe to the innateness hypothesis view language as a skill or system of conventions that has been developed and learnt merely on the basis of domain-specific cognitive capacities. The task that evolutionary linguists face if they attempt to account for the origins of language consequently depends on the conception of language they start from: depending on their position on the innateness question, they have to explain the emergence of a biological entity or the emergence of a cultural entity.

1.1.1 Language as a biological entity

The most seminal nativist approach to language evolution was proposed by Pinker and Bloom (1990). Pinker and Bloom combine the innateness hypothesis with the observation that language exhibits complex adaptive design for its function in communication. They conclude that human language must be the product of natural selection since natural selection is the only process that is capable of accommodating both premises. With their proposals, Pinker and Bloom (1990), and later Jackendoff (2002) and Pinker (2003:21–27), respond to nativist accounts of language evolution that do not consider language to be adaptive (specifically Chomsky 1988; Piatelli-Palmarini 1989; Gould 1997), as well as to selectionist theories which attempt to explain language as an adaptation to something other than communication (Dunbar 1998; Miller 2000). The latter group represents approaches which are typically only marginally informed by linguistics. Such theories, which invoke putative selection pressures such as grooming (Dunbar 1998) or courtship (Miller 2000), are flawed since, even though they refer to natural selection, they cannot account for the particular design that language has.

The cornerstone of all nativist theories of language evolution is the innateness hypothesis, the claim that the core of language is genetically determined in a language-specific cognitive module. The innateness assumption is mainly based on three arguments: (i) the argument from the poverty of the stimulus, (ii) the argument from the universality of language, and (iii) the argument from linguistic universals. The argument from the poverty of the stimulus, as put forward by Chomsky (e.g. 1959, 1965, 1981) and, more recently, Laurence and Mangolis (2001), roughly states that the linguistic data that is available to children is insufficient for them to acquire language by means of grammar induction, and that it must therefore be concluded that children possess an innate blueprint for language (usually referred to as "language acquisition device," "Universal Grammar," or "narrow faculty of language"), which reduces the task of language acquisition to a mere setting of culture-specific parameters within a genetically determined computational system. The argument from the poverty of the stimulus has been contested on various grounds. It has been claimed that there is no poverty of the stimulus if one presupposes a model of learning that is more realistic than the sort of isolated grammar induction that the argument is built on (e.g. Tomasello 1995, 2003a). Others have pointed out that the input that children receive is actually not as impoverished as the argument from the poverty of the stimulus presupposes (e.g. Marcus 1993; Saxton 1997; Saxton et al. 2005; Tellier 1999, 2000; Chouinard and Clark 2003). Similar cases against the argument from the poverty of the stimulus have been made by Arbib and Hill (1988) and Elman et al. (1996). Zuidema (2003) finally turns the argument on its head and suggests that children are able to learn language because language can only evolve structures that are learnable in the first place: "[t]he poverty of the stimulus solves the poverty of the stimulus."

While the argument from the poverty of the stimulus remains the main argument for the innateness hypothesis—most of the arguments listed in Pinker (2003:22f.) can be viewed as aspects of the poverty of the stimulus argument—two minor additional arguments deserve some brief mentioning at this point. The first concerns the universality of language. All human societies without exception possess language. Nativists take this fact as a point in favour of the innateness hypothesis. Language, the claim goes, is different from "unambiguously culturally acquired abilities" (Pinker 2003:22) such as agriculture, chess, mathematics, or government. Its universality implies that it is rather on par with "instincts" such as fear or sexual desire. This argument, of course, can easily be refuted. If one assumes that (i) all humans possess the same general cognitive and social capacities, and that (ii) all humans face the same basic communication tasks, it is hardly surprising that they have all developed some sort of language—just like toolmaking is universal too. Another argument is that from linguistic universals. All languages exhibit some general characteristics, sometimes called the design features of language (Hockett 1960; Greenberg 1963; Comrie 1981). There are two problems with invoking this fact to make a case for the innateness hypothesis. Firstly, there is no reason, on the basis of the observation alone, to assume that linguistic universals reflect a language-specific Universal Grammar rather than properties of general cognitive capacities. Secondly, there is immense disagreement within the nativist camp itself on what actually constitutes Universal Grammar. Most recently, Hauser et al. (2002), in retreat from earlier, more extensive claims, postulated that recursion is the only (and therefore last remaining) component of Universal Grammar. Parker (2006), however, argues that recursion is neither specific to humans nor specific to language and can therefore, by definition, not constitute a part of Universal Grammar.

At the present stage, the question of whether some part of language must be regarded as a domain-specific biological entity remains unresolved: the innateness hypothesis has neither been sufficiently corroborated nor disproved. Pullum and Scholz (2002), and more recently Scholz and Pullum (2006), come to the conclusion that the advocates of the poverty of the stimulus argument have, as yet, not produced sufficient evidence to motivate the innateness assumption. Tomasello (2004) even goes as far as to claim that the innateness assumption must be rejected as a scientific hypothesis altogether because it cannot be falsified. The present thesis does not aim to make a case either for or against the innateness hypothesis. It will, however, make a methodological decision. It is good scientific practice to assume as few assumptions as possible and to make an assumption only if some phenomenon could not have been explained without it. It seems to me that the explanatory power of domain-general cognitive capacities with regard to explaining language have so far not been exhaustively investigated. My thesis a will thus, for methodological reasons, not make the innateness assumption. That is, I will treat language as a cultural rather than a biological entity.

1.1.2 Language as a cultural entity

Non-nativist approaches to language and language evolution (e.g. Deacon 1997; Tomasello 1999) do not entertain the idea that there is a domain-specific language "organ" but rather view language as a cultural entity that has been learnt and developed purely on the basis of domain-general human cognitive capacities. From this perspective, the study of language evolution has two aims: (i) to identify the minimal set of general cognitive capacities that is necessary for language to emerge, and (ii) to account for what Kirby (1999:20) calls "the problem of linkage," namely to explain by means of what processes one gets from this set of domain-general cognitive capacities to language as we know it today. The mechanisms by which activities grounded in general human intelligence result in the emergence and further development of cultural entities (skills, tools, social institutions) are commonly summarised under the label "cultural evolution." The term has come to include the component of "evolution" because it has been observed that, on a very abstract level, the processes it describes resemble those involved in biological evolution (Hull 1988; Croft 2000). However, as Pinker (2003:21) points out, for non-nativist explanations of language evolution to be on par with biological explanations, it does not suffice to simply refer to general intelligence and cultural evolution: what is needed is rather the development of mechanistic models that describe the respective processes. In recent years, this requirement has begun to be met by a growing body of conceptual, mathematical and computational models of the cultural evolution of language (e.g. Brighton et al. 2005; Christiansen et al. 2002; Hurford 2002; Kirby 2000; Kirby and Hurford 2002; Smith 2006b).

Most of these models focus on the role that (potentially imperfect) learning plays in the emergence of linguistic features (see Smith 2006b for an overview). However, this focus on learning as a driving-force in the cultural evolution of language restricts the explanatory power of the respective models with regard to the two evolutionary puzzles introduced at the beginning of this chapter: the emergence puzzle and the design puzzle. First, the fact that such models typically work with more or less language-specific learning algorithms entails that they need to assume some sort of language as their starting point, usually a hypothesised protolanguage that possesses symbolism but no or only very little grammatical structure. These models thus do not explain how language has emerged from no language but only how a simple form of language can turn into a more complex one.[1] Second, while they account for why certain linguistic features seem to be particularly well adapted to being learnt, they do not explain how

[1]Note that the fact that they deal with a qualitative and quantitative leap in the evolution of language distinguishes such models from mere models of language change, which only describe the transition from one state of language to another, *equally* complex state of language.

language comes to exhibit the appearance of design for communication. This thesis addresses these two shortcomings of existing models by complementing them with a model that focuses on the role of language use rather than language learning as a driving-force of linguistic cultural evolution.

1.2 Pragmatic plasticity

The aim of this thesis is to develop a mechanistic model of the cultural evolution of language that focuses on the role of language use. In contrast, most existing mechanistic models study the role of language learning. To isolate the effects of language learning, they abstract away from the mechanisms of use. Specifically, they assume that linguistic signals fully specify the meaning they communicate and ignore the fact that this is simply not the case in real language use. The basic distinction that models with such an idealised representation of language use fail to incorporate is that between *signal meaning* and *speaker meaning* (e.g. Grice 1957, 1968; Clark 1996). Speaker meaning refers to the information that the speaker actually intends to communicate in a particular communicative act. Signal meaning, on the other side, denotes the meaning that is conventionally associated with the produced linguistic form in the user's linguistic knowledge, the meaning that is encoded in a linguistic signal.[2] That the two are not identical was first pointed out by Grice (1957). Language exhibits what I will call *pragmatic plasticity*: the speaker meaning that a signal meaning develops into is determined

[2]There is some disagreement among pragmaticists over the nature of signal meaning and speaker meaning, and over the question of whether an intermediate level of meaning needs to be postulated and different inferential processes have to be distinguished (see e.g. Huang 2007:209–244 and Carston 2002:94–221 for overviews). In this thesis, I will understand the two concepts in the broad, relevance-theoretic sense outlined in Wedgwood (2007). Relevance Theory (Sperber and Wilson 1995; Carston 2002) recognises "just two theoretically and psychologically levels of meaning," namely (i) "that which is communicated by a given linguistic form independently of context," and (ii) "that which is understood to be communicated in a particular context—which is an indeterminate number of propositions, among which implicatures and any sense of 'what is said' are derived by the same mechanisms" (Wedgwood 2007:678). Signal meaning therefore is neither assumed to be truth-conditional nor fully propositional (in fact, in most of the following examples, the signal meaning is sub-propositional), and no psychologically relevant distinction between different types of inference from context is postulated. Like Relevance Theory, I start out "from a much more basic position on what kinds of meaning ('content') there must be: simply that some meaning is encoded in linguistic forms and some is inferred" (Wedgwood 2007:652).

by its context of use. Signal meanings become speaker meanings when they are interpreted in specific contexts. [3]

1.2.1 Underspecification

During the process of interpretation in context, signal meanings are regularly complemented with additional, pragmatically inferred information. In such cases, the signal meaning *underspecifies* the speaker meaning: the meaning encoded in the signal contains less information, is less specific, than the actually communicated meaning. The following examples illustrate the ubiquity of underspecification in language use.

Pragmatically inferred information may be used for *disambiguation* in cases where the produced linguistic signal is ambiguous, that is, where it is conventionally associated with more than one meaning. Example (1) represents a case of lexical ambiguity: the word *port* can stand for either fortified wine or a harbour. The hearer therefore has to infer additional information from the context, namely which of these two meanings the speaker refers to. Similarly, the signal meaning also underspecifies the speaker meaning in cases of syntactic ambiguity (2), scope ambiguity (3), or plural ambiguity (4).

(1) a. John and Bill passed the port.
    b. John and Bill passed the port [wine].
(2) a. Only old men and women live in this building.
    b. Only old men and [old] women live in this building.
(3) a. Every linguist speaks two foreign languages.
    b. Every linguist speaks [the same] two foreign languages.
(4) a. Two men lift a table.
    b. Two men lift a table [each].

The signal meaning also underspecifies the speaker meaning in the sense that it is in need of *reference resolution*. Reference resolution subsumes anaphora resolution as well as symbol grounding. The interpretation of sentence (5), for instance,

[3] The term "pragmatic plasticity" bears a partly deliberate resemblance to the concept of "developmental plasticity" in biology. Just as the conventions of a linguistic code only partly specify the meaning an utterance is going to convey (the remainder being influenced by the context, due to pragmatic plasticity), the information in the genotype of an individual only partly specifies the properties of its phenotype (the remainder being influenced by the environment, due to developmental plasticity). This, however, is as far as we should take the metaphor; and it must be emphasised that it is just that: a metaphor.

includes among other things working out which real persons the names *Elizabeth* and *Susan* designate as well as which of the two the pronoun *her* refers to. A further case where reference resolution needs to be performed is deictic expressions like the one in example (6): the information that the demonstrative pronoun *this* refers to a specific crack in the window has to be inferred from the context.

(5) a. Elizabeth asked Susan to pass her the salt.

b. Elizabeth$_{[i][Elizabeth']}$ asked Susan$_{[Susan']}$ to pass her$_{[i]}$ the salt.

(6) a. Look at this!

b. Look at this$_{[crack']}$!

Saturation has been described as the process whereby a variable or slot opened up by the linguistically encoded meaning is filled (Recanati 1989, 2004a). Prominent examples of saturation are the specification of elliptical constructions (7) or of vague relations such as the one between possessor and possession in possessive constructions (8) or between two compound nouns (9). Here too, the signal meaning underspecifies the speaker meaning.

(7) a. Joan is cleverer.

b. Joan is cleverer [than Naomi].

(8) a. I enjoyed reading John's book.

b. I enjoyed reading the book [written by] John.

(9) a. She now works for the Boston office.

b. She now works for the office [located in] Boston.

Free enrichment is similar to saturation.[4] It describes situations where, "although there does not seem to be either an overt indexical or a covert slot in the linguistically decoded logical form of the sentence uttered, the logical form nevertheless needs to be conceptually enriched" (Huang 2007:191). Examples (10)–(12) show such cases of signal meanings conceptually underspecifying their respective speaker meanings.

(10) a. John has a brain.

b. John has a [scientific] brain.

[4]The distinction between "saturation" and "free enrichment" is Recanati's (e.g. 2004a; 2004b). The relevance theorists subsume both under the notion of "enrichment" (e.g. Sperber and Wilson 1995; Carston 1997).

(11) a. The police moved in and the hostages were released.

b. The police moved in and [then, as a result] the hostages were released.

(12) a. I have nothing to wear.

b. I have nothing [suitable] to wear [to John's wedding].

1.2.2 Overspecification

The process of interpretation in context also often leads to some aspects of the signal meaning being discarded. In such cases, the signal meaning *overspecifies* the speaker meaning: the meaning encoded in the signal contains more information, is more specific, than the actually communicated meaning.

Like underspecification, overspecification occurs in a vast variety of linguistic phenomena. As, for instance, Sperber and Wilson (1995:233–37) and Langacker (1987:70) point out, overspecification is particularly prominent in figurative language use.[5] Sperber and Wilson (1997:105) find fault in the fact that figurative language use has traditionally been investigated in isolation from other pragmatic phenomena, each having produced their own literature. "There is little overlap between the two, as if it went without saying that these are two essentially different aspects of language." They locate the reason for this development in the fact that "[i]n both cases, the literature is centrally concerned with problems of classification and offers little in the way of explanation" (Sperber and Wilson 1997:105). They claim to offer a more integrated approach, which is further elaborated especially by Carston (1997) and, recently, Wilson and Carston (2007). These studies treat what Carston (1997) refers to as "enrichment" and "loosening" in a largely symmetrical way, as two sides of the same coin: consequences of underspecification and overspecification respectively. The following examples illustrate some of the linguistic phenomena that are based on overspecification.

The signal meaning overspecifies the speaker meaning in cases of *approximation*, that is, where an expression is used to convey a speaker meaning that is less restrictive than what the expression actually stands for (see e.g. Wilson and Carston 2007). In such examples, certain specifications contained in the signal meaning are ignored. In example (13), for instance, France is described as a hexagon, even though, strictly speaking, it is only vaguely similar to one. The notion of *hyperbole*

[5] Sperber and Wilson (1995:233ff.) refer to the use of overspecified expressions as "loose talk"; Langacker (1987:69ff.) calls it "partial sanction."

or overstatement is closely related. It is exemplified in sentence (14). One aspect of bankruptcy is the loss of a lot of one's money. This part of the encoded meaning is kept whereas the other elements of it, for instance the aspect of ending up with no money anymore at all, are discarded.

(13) a. France is a hexagon.
     b. France has roughly the shape of a hexagon.

(14) a. This policy will bankrupt the farmers.
     b. This policy will make the farmers lose a lot of their money.

In both approximation and hyperbole, some aspect of the encoded meaning is cancelled during utterance interpretation. Overspecification is even more prominent in *metaphors*. In example (15), it is only very few of the various features that make up a chameleon which are kept, namely the ones that chameleons and Sally have in common. In this case, it is probably the property of easily adapting to one's environment or the property of frequently changing one's appearance. Other features of chameleons, for example their property of having a long tongue or the fact that they are reptiles, are ignored because they evidently do not contribute to what the speaker intends to communicate (unless, of course, Sally lived in a zoo and were indeed a chameleon).

(15) a. Sally is a chameleon.
     b. Sally frequently changes her appearance.

Overspecification is also present in cases of *metonymy*. Two prominent features of the White House can be described as follows: the White House is (i) a building, and (ii) an entity related to the office of the president of the United States. In example (16), the former aspect of the signal meaning of *the White House* is ignored and only the latter contributes to the speaker meaning.

(16) a. The White House issued a statement.
     b. (An entity related to) the office of the president of the United States issued a statement.

These are just a couple of examples that illustrate the ubiquity of underspecification and overspecification in language use. It goes without saying that the two aspects of pragmatic plasticity can, and in fact most often do, occur in combination: in most, if not all, instances of language use, the signal meaning is (i) complemented with additional, pragmatically inferred meaning while, at the same time, (ii) some of the information contained in the signal meaning itself is discarded and does not contribute to the speaker meaning. The notion of pragmatic plasticity—that is, the fact that signal meanings under- and/or overspecify the speaker meaning into which they develop in specific contexts—and its explanatory potential with regard to language evolution will be the focus of this thesis.[6]

1.3 Outline of the thesis

In the last part of this introduction, I provide a thesis outline comprised of three parts. I will first specify objectives and scope of the thesis. I will then briefly mention the main methodological commitments that I make in order to address the inherent lack of data that any study of language evolution naturally faces. Finally, an overview over the organisation of the thesis and the content of its individual chapters will be given.

1.3.1 Objectives

The aim of this thesis is (i) to develop a mechanistic model of the cultural evolution of language that acknowledges and incorporates the fact that language use exhibits pragmatic plasticity and (ii) to explore the explanatory potential of this fact with regard to the two evolutionary puzzles introduced at the beginning of this chapter:

- *The emergence puzzle*
  Language has emerged from no language.

- *The design puzzle*
  Language exhibits the appearance of design for communication.

The model to be developed aims to explain the evolution of language on the basis of domain-general cognitive capacities and mechanisms and does not make

[6]I prefer the term "pragmatic plasticity" over the more common term "linguistic underdeterminacy" because the former simply reflects the fact that speaker meanings can deviate from signal meanings whereas the latter appears to emphasise one possible form of deviation (underspecification) more than the other (overspecification).

the innateness assumption. Some recent models of the cultural evolution of language have started to look at the effects that the emergence of culture, and in particular language, may have on the subsequent course of biological evolution, and the interaction between biological and cultural evolution (e.g. Kirby et al. 2007). Such considerations are outside the scope of this thesis.

Finally, I aim to describe the model at a level of abstraction that allows for it to be integrated with both models of general, non-linguistic cultural evolution and models of historical language change. To maintain sufficient specificity and formal rigour despite such a relatively high level of abstraction, I formulate the model conceptually and also implement it computationally.

1.3.2 Methodological commitments

Evolutionary linguistics is a field that is known for its notorious lack of evidence (Cangelosi and Parisi 2002a:4). Its object of study, the origin of language, is simply too far remote in the pre-historic past for any direct observation to be available. Under such circumstances, methodological considerations attain particular significance. It is therefore appropriate at this point to state some broad methodological commitments underlying this thesis. The first commitment falls within the range of what is known as *Ockham's razor*: entia non sunt multiplicanda praeter necessitatem. If multiple explanations for a phenomenon are possible and all other aspects are equal, the explanation which introduces fewer assumptions and postulates fewer hypothetical entities is to be preferred. I am aware that this principle has led to rather different interpretations, particularly within linguistics, and has thus at the same time given rise to theories as diverse as, for instance, the Minimalist Program (Chomsky 1995b) and Cognitive Grammar (Langacker 1987). However, this fact does not diminish the general methodological value of the principle.

The second commitment is that of *uniformity of process*. This principle too has been subject to various different linguistic interpretations (see overviews e.g. in Christy 1983; Newmeyer 2003). I take it to be applicable to evolutionary linguistics in the following form: "the same mechanisms which operated to produce the large-scale changes of the past may be observed operating in the current changes taking place around us" (Labov 1982:161, cited in Trask 2000:354). The consequence of this interpretation of the principle is that the processes invoked in a model of language evolution must be attested either in present day language use, language change, or language acquisition. I will argue in chapter 2 that, contrary to what is implied by current acquisition-oriented models of the cultural evolution of language, it is the first two (language use and change) which are of particular relevance to evolutionary linguistics. Wherever possible, assumed processes must thus be corroborated by attested linguistic evidence.

1.3.3 Organisation

This thesis comprises two parts. In the first part (chapters 2 and 3), I develop the model conceptually and work towards an account of the emergence puzzle. In the second part (chapters 4 and 5), I propose a computational implementation of the introduced conceptual model, and, with the help of this, I address the design puzzle. The individual chapters are organised as follows:

Chapter 2. I begin by analysing the components of extant models of general cultural evolution and discuss how models of language change and linguistic evolution map onto them. I identify innovative use as the motor of cultural evolution.

Chapter 3. I then conceptualise the cognitive mechanisms underlying innovative use and argue that they originate in pre-linguistic forms of ostensive-inferential communication. In order to account for the emergence puzzle, I use the identified mechanisms to provide a unified explanation of the emergence of symbolism and the emergence of grammar. Finally, I discuss the implications of the presented analysis for the so-called proto-language debate.

Chapter 4. I then propose a computational implementation of the developed conceptual model. This computational implementation allows for the simulation of the cultural emergence and evolution of symbolic communication and provides a laboratory-like environment to study individual aspects of this process.

Chapter 5. I address the design puzzle by employing such computer simulations to explore the role that pragmatic plasticity plays in the development of the expressivity, signal economy and ambiguity of emerging and evolving symbolic communication systems.

Chapter 6. In the conclusion of the thesis, I re-assess the explanatory potential that the introduced model—and in particular the fact that language use exhibits pragmatic plasticity—has with regard to the study of the origins of language and point to avenues of future research.

CHAPTER 2

The cultural evolution of language

Nativists have frequently criticised non-nativist approaches by asserting that there is a linkage problem between the properties of general intelligence and the structure of language.[1] Many linguistic peculiarities simply do not fall out of what we know about human cognition or functional needs. Cognitive capacities do not, for an overwhelming number of cases, directly map onto linguistic features. If one intends to challenge the innateness hypothesis, one therefore has to find a way of bridging this gap: how do we get from human cognition to the structure of language?

The missing link, it has been suggested, is cumulative cultural evolution (e.g. Hurford 1990; Kirby 1999; Tomasello 1999). Language is continuously learnt and modified, and the innovations added by speakers from one generation build upon those made by earlier generations. But cultural evolution is, of course, not limited to language but can be applied to wide varieties of artefacts and social practices. And while there seems to be some agreement that cultural evolution has the capacities to bridge the described gap, a whole plethora of processes and mechanisms has been suggested to describe the exact means by which cultural

[1] The criticism raised by nativists and the non-nativist response is reflected in the row that arose over the publication of Newmeyer (1991). In this article, Newmeyer alleges that "in the functionalist view, grammatical patterning mirrors discourse function in a *direct* way" (Newmeyer 1991:4, emphasis added). This statement, and the article that contained it, has consequently attracted a lot of negative reactions (among others Hopper 1991; Lakoff 1991; Mühlhäusler 1991). A reaction to the particular statement above comes from Thompson (1991:93):

> [H]e claims that "... in the functionalist view, grammatical patterning mirrors discourse in a direct way." It is noteworthy that Newmeyer does not cite any references here, and indeed, I believe he would have a difficult time doing so, for I am not aware of any functionalist who would take such a position.

evolution in general and the cultural evolution of language in particular comes about.

In the first part of this chapter, I introduce cultural evolution from a general, anthropological point of view and discuss the main building-blocks of a model of cultural evolution, and the controversies surrounding them (section 2.1). I will then move on to ask how such a general model of cultural evolution can be applied to language. Two of the questions that fall out of this enterprise will be tackled in this chapter, namely (i) what the linguistic artefact is (section 2.2) and (ii) where the main locus of linguistic innovation lies (section 2.3). A third question, the one about the function and usage of the linguistic artefact, will be addressed in the next chapter. The aim of this chapter is to show what aspects a *general* model of the cultural evolution has, why existing models of the cultural evolution of *language* neglect some substantial mechanisms, and how they can be enhanced to provide a more complete picture of what is going on during the evolution of language.

2.1 Cultural evolution

Cultural evolution is a concept which has mainly been developed and applied in anthropology, archaeology and philosophy, and before I engage in explaining its central ideas and concepts, a brief overview of the most influential studies contributing to this new set of theories seems to be appropriate. Dawkins (1979) suggested that the evolution of culture may function in ways which are fundamentally similar to those of biological evolution. In their seminal book, the anthropologists Boyd and Richerson (1985) take this idea up and show how human culture can be described as an evolutionary system. They develop mathematical models to study how processes of this system work and how they interact with biological evolution. Another seminal paper in the recent history of the study of cultural evolution is Tomasello et al. (1993), which points out the cumulative nature of human culture (as opposed to cultures in other animals), that is, the fact that "none of the most complex human artifacts or social practices—including tool industries, symbolic communication, and social institutions—were invented once and for all at a single moment by any one individual or group of individuals" (Tomasello 1999:5). Tomasello et al. (1993) coin the term "ratchet-effect" for this phenomenon. Boyd and Richerson (1996) propose an explanation for why cumulative culture is rare in animals other than humans. Tomasello (1999) spells out the ratchet-model of cultural evolution and applies it to human language acquisition. Its main points are summarised in later papers such as, for example, Tomasello (2002) and further elaborated with respect to language acquisition in Tomasello (2003a). I provide a detailled description of Tomasello's ratchet model of cumulative cultural evolution later in this chapter.

The idea of cultural evolution has also entered theories of archaeology, especially under the influence of the so-called processural archaeology, which aims at going beyond a mere cataloguing of excavated artifacts and at gaining insights into the general processes at work in cultural change. While anthropologists tend to refer to these processes as "cultural evolution," the term "Darwinian archaeology" is more commonly used in the field of archaeology. Strictly speaking however, as Shennan (2002:15) in reference to Smith (2000) points out, Darwinian archaeology comprises "three styles in the evolutionary analysis of human behaviour" (evolutionary psychology, behavioural ecology, and dual inheritance theory) only one of which (dual inheritance theory) is equivalent to cultural evolution. O'Brien and Lyman (2000, 2002) discuss the applicability of theories of cultural evolution to archaelogy from a methodological point of view and provide a concise overview of the current state of the field.

Other noteworthy work by philosophers, psychologists and biologists who apply typical concepts of biological evolution to cultural processes and thereby attempt to establish similarities and differences between the two domains includes, for example, Hull (1988); Dennett (1995); Sperber (1996); Odling-Smee et al. (2003); Sterelny (2006). This is just to mention some of the most important and accessible studies of a fast growing interdisciplinary field. I will now proceed to explaining the concept of cultural evolution in more detail.

2.1.1 What is culture?

Culture is usually defined as behaviour which is neither determined genetically nor learnt individually but which is acquired through some form of social learning. It is therefore one of three possible ways for an individual to acquire knowledge, and in order to understand its evolutionary role it is useful to look at the relation of all three—*genetic encoding, individual learning*, and *social learning (culture)*—to each other. In the following paragraphs, I will therefore briefly outline the differences between genes and individual learning, and between individual learning and social learning respectively, and explain how culture combines the advantages of both other forms of knowledge acquisition to which evolution has given rise.

2.1.1.1 *Genes vs individual learning*

The behaviour of an individual can be determined, as the result of natural selection, by its genes. Due to that behaviour, certain individuals, whose genes elicit some novel behaviour, may have a greater chance to survive and pass on their genetic information in some particular environment than their conspecifics whose genes do not give rise to it. Individual learning, on the other hand, denotes behaviour which is not genetically encoded but acquired by one individual in a specific environment via trial-and-error. A chimpanzee who figures out how to fish for termites in a mound by using a small stick, for instance, does not exhibit genetically encoded behaviour but individual learning.[2] In this case, learning can be described as the solving of a problem imposed on the chimpanzee by the environment.

The advantage of the capacity for individual learning over exclusively genetically determined behaviour is an increased flexibility of behaviour, which brings about better adaptability to the environment. Such behavioural plasticity, that is, the flexibility with which an individual can react to and interact with its environment, is one way in which evolution can increase that individual's fitness. Individual learning can be especially advantageous in a rapidly changing environment. Evidently, the most impressive example in this context is provided by humans. The extraordinary capacity for individual learning exhibited by humans has contributed to their spreading over the planet into environments as diverse as Greenland and the Sahara desert. It has to be noted at this point though that, as we shall see below, individual learning can only do the job if its outcome is transmitted faithfully among conspecifics.

2.1.1.2 *Individual learning vs culture*

Social learning, as opposed to individual learning, is usually defined as any form of learning the outcome of which is due to the outcome of the learning of a conspecific. Shennan (2002:37) exemplifies the distinction between individual learning and social learning by stating that

[2] Note, however, that the general *disposition* for individual learning is of course again determined genetically.

> [n]ot all information capable of affecting people's phenotypes is cultural, only that acquired from other members of the same species. Thus, if I learn to make a stone tool by experimenting with striking pieces of stone together, what I learn does not count as cultural information. Rather, it is the sort of learning from interacting with the environment that is characteristic of many, if not most, animals. However, if my daughter learns by watching me then the practices she acquires in doing so count as cultural.

The advantage of social learning (culture) over individual learning is that inheriting information from conspecifics saves the cost (time, effort, risk) one would have to invest if one had to make the discoveries by oneself through trial and error. Rat pups, for instance, avoid the unnecessary risk of consuming potentially dangerous food by only eating substances which they can identify in the milk of their mothers or in the odours of conspecifics (Galef 1996). Similarly, a chimpanzee who is looking for food and accidentally finds some edible insects under a log involuntarily allows conspecifics who were observing to access the food source directly without having to search for it themselves. Probably the most eminent example of saving time and effort through social learning is human instructive learning. It is much easier if somebody tells you how to decipher Morse code or to drive a car, or simply how to peel an avocado, than if you had to figure it out by yourself. Tomasello (1999:4) summarises this by stating that

> [b]roadly speaking, cultural transmission is a moderately common evolutionary process that enables individual organisms to save much time and effort, not to mention risk, by exploiting the already existing knowledge and skills of conspecifics.

Another advantage of some forms of social learning (e.g. the aforementioned instructive learning) over individual learning is that they allow for cumulative change—a concept to which I will turn soon.

2.1.1.3 Genes vs culture

Individuals can inherit information or behaviour from their conspecifics in two ways, namely either through their genes or via cultural transmission (social learning). Avital and Jablonka (2000:54, cited in Shennan 2002:36) define such inheritance systems as "the regeneration of phenotypic traits and processes through the direct or indirect transmission of information between entities," and Shennan (2002:16) summarises the relation between culture and genes as follows:

> [V]ariation in human behaviour cannot be explained solely in terms of criteria linked to reproductive success. Humans have a second inheritance system, culture, in addition to their genes. Cultural traditions are handed down from one generation, and indeed from one day, to the next, by specifically cultural mechanisms.

The advantage of cultural inheritance over genetic inheritance is that culture can react to a dynamic environment on much smaller time-scales than biological evolution. This is due to two reasons. First, cultural innovations are, at least to a certain extent, purposive whereas genetic mutation is not. Mutations are random "copy errors" and, metaphorically speaking, this renders biological evolution in the position of a blind person searching for the exit of a room. Cultural innovations, on the other side, can be more target-oriented: innovations through individual learning are usually produced as reactions to specific novel situations that have been encountered. Second, innovation can happen more frequently in culture than in biological evolution. Genetic mutation happens once per generation, or rather from one generation to the next in biological evolution. The equivalent cultural processes can occur with much shorter intervals, "from one day to the next" (Shennan 2002:16). Cultural innovations are made not just when knowledge is passed on to the next generation but can happen whenever that knowledge is applied in some particular situation. This is a crucial point to which I will return later in this section. Culture thus combines the advantages of both learning and inheritance: it flexibly adapts to its environment through learning but is also economical since it is inherited via social transmission.

2.1.2 What is cultural evolution?

The term "cultural evolution" is usually applied to the set of processes through which *cumulative culture* (as opposed to *non-cumulative culture*) comes about. Tomasello (e.g. 1999:5) tends to speak of "cumulative cultural evolution." In this thesis, I distinguish between the process (cultural evolution) and the product (cumulative culture) where such a distinction is helpful, and use the two terms interchangeably otherwise. The fact that "evolution" has been adopted as an element of the term to describe the processes leading to cumulative culture points to the assumption that these processes share some underlying properties with the ones at work in biological evolution (Shennan 2005:49). Thus, one has to distinguish between cumulative culture and non-cumulative culture. This differentiation has been introduced by Tomasello et al. (1993), who coined the term "ratchet effect" for the main property that leads to cumulative culture, and has been developed further in Boyd and Richerson (1985, 1996); Tomasello (1999, 2003a). In particular Boyd and Richerson (1996) and Tomasello (2002) discuss what separates these two forms of culture. The nature of cumulative culture is therefore best explained if it is considered in contrast to non-cumulative culture.

2.1.2.1 Non-cumulative culture

The main characteristic of non-cumulative culture is that cultural traits do not reach a state where they could not be invented entirely by any one individual on its own. "With few exceptions, social learning leads to the spread of behaviors that individuals could have learned on their own" (Boyd and Richerson 1996:5). Tomasello (1999) consequently distinguishes between weak and strong forms of social learning and claims that only the latter can create cumulative culture. As examples of weak forms of social learning, leading to non-cumulative culture, he discusses local enhancement, emulation learning, and ontogenetic ritualisation. Strong forms of social learning (that is, the "few exceptions" mentioned by Boyd and Richerson above), on the other side, seem to be limited to imitative learning (sometimes also called "observational learning") and active instruction.

> Tomasello et al. (1993) argued that cumulative evolution depends on imitative learning, and perhaps active instruction on the parts of adults, and cannot be brought about by means of "weaker" forms of social learning such as local enhancement, emulation learning, ontogenetic ritualisation, or any form of individual learning. (Tomasello 1999:39)

The fact that weaker forms of social learning cannot create cumulative culture may be due to two reasons. In a first group of cases, the involved form of social transmission does not actually replace individual learning but just facilitates and triggers it. This becomes clear from the discussion of local enhancement, emulation learning and ontogenetic ritualisation offered by Tomasello (1999:26–33).

Local enhancement "occurs when the activity of older animals increases the chance that younger animals will learn the behaviour on their own" (Boyd and Richerson 1996:7). As an example for local enhancement, Tomasello (1999:26–28) mentions the potato washing in some groups of Japanese macaques documented by Kawamura (1959); Kawai (1965). The story goes like this. At some point, a macaque figured out that the sweet potato it intended to eat could be freed of the sand sticking to it if it was washed in water. The fact that the offspring of the inventor followed their mother around—and therefore also to the nearby water where she used to wash her potatoes—increased the likelihood that they discovered the technique as well. Like this, the cultural trait spread in the whole local population without any one macaque actually copying the behaviour of a conspecific.

Emulation learning describes cases where an individual is given the prerequisites for its own learning by the way the learning of another individual has changed the environment. Emulation learning is given, for instance, when a chimpanzee watches a conspecific who turns a log to get to some edible insects hidden underneath (Tomasello 1999:29). While the observer will not imitate the observee, it still knows, because of the observee's behaviour, where the insects can be found and will thus sooner or later start to figure out on its own that it can reach the desired source of food by turning around the said log. Again, the individual learning of the first chimpanzee has merely enhanced the individual learning of the second, or of the whole group eventually.

"In *ontogenetic ritualization* a communicatory signal is created by two organisms shaping each other's behavior in repeated instances of a social interaction" (Tomasello 1999:31, emphasis added). If a young chimpanzee wants to climb its mother, it will start doing so by lifting its arms towards the adult's body. If the youngster's arm-lifting repeatedly initiates its climbing the mother, the adult will, at some point, predict the young chimpanzee's intention just from that first initiating gesture and lift it up herself. This behaviour of the mother will in turn make the youngster realise that it is apparently enough to lift one's arms in order to be lifted by the mother. It will therefore not try to execute the whole climbing procedure anymore but content itself with the initial gesture to signal to the adult that it wants to be lifted up. Mother and child have thus created some form of "private culture." Their behaviour could not have been acquired through individual learning by just one party, and neither is it encoded genetically. In all three cases, local enhancement, emulation learning, and ontogenetic ritualisation alike, the actual discovery has thus to be made by each individual anew—none of them *directly* builds on the achievements of conspecifics.

In a second group of cases, the involved form of social transmission is encoded genetically so that it is limited to very specific cases of learning and cannot be extended flexibly to novel situations or learning tasks. Examples for this type of culture can be found in ants and rats. Ants locate food by following the pheromone trails of conspecifics, and thus save the time, energy and risk of searching for it by themselves (Tomasello 1999:4, referring to Mundinger 1980; Galef 1992). Rats only eat food they can identify in the milk of their mother or the odour of conspecifics. The food preferences of one rat can thus spread in a population via its offspring, as mentioned in Tomasello (1999:4) and Boyd and Richerson (1996:5, referring to Galef 1988). In both cases, the rats' and the ants', social transmission is limited to one very specific genetically encoded behaviour and can thus not give rise to an accumulation of novel behavioural strategies on a cultural rather than biological time-scale.

In all these cases, social transmission only enhances individual learning but does not replace it. The output of the learning of conspecifics might trigger or facilitate the learning of an individual, but the individual usually still has to go through the process of discovery by itself, or does at least not acquire any behaviour which it could not have "invented" on its own.

2.1.2.2 Cumulative culture

Traits of cumulative culture, on the other hand, are not the result of one single incident of individual learning but rather represent the accumulation of many modifications made to an original invention by a whole number of individuals over time. Tomasello (1999, 2003a) has formulated an abstract model of cultural evolution which explains how such cumulative culture emerges. He describes the main idea of his model of cultural evolution as follows:

> [S]ome individual or group of individuals first invented a primitive version of [an] artifact or practice, and then some later user or users made a modification, an "improvement," that others then adopted perhaps without change for many generations, at which point some other individual or group of individuals made another modification,

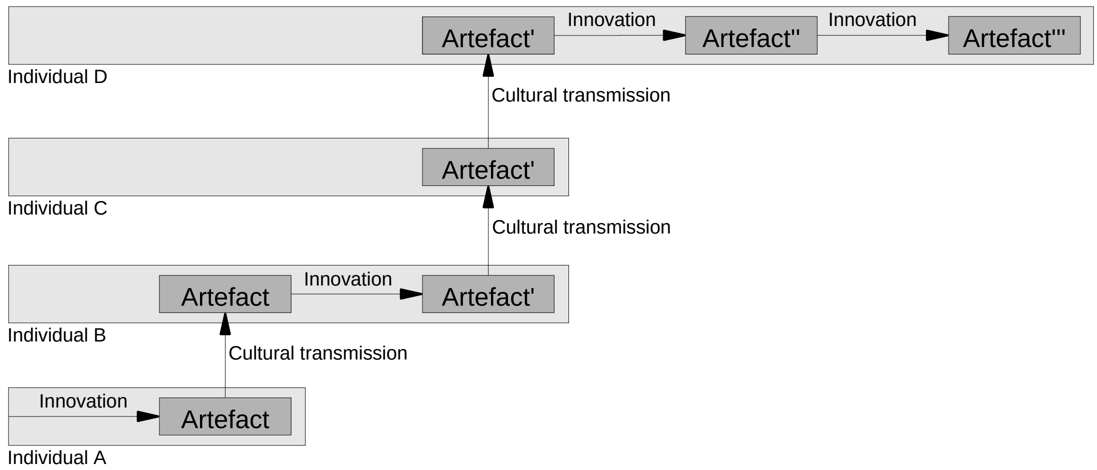


Figure 2.1: The ratchet effect in cumulative culture. In a given environment, individual *A* invents an artefact to perform a desired function. The knowledge of how to make and use the artefact is then passed on from individual *A* to individual *B*. Individual *B* modifies the artefact, maybe to make it perform its function in a novel environment, before it passes its knowledge on to individual *C*. Individual *C* uses the artefact but does not modify it. The artefact is thus passed on to individual *D* in the state in which it had been inherited from individual *B*. Individual *D* adds two modifications to the artefact.

> which was then learned and used by others, and so on over historical time in what has sometimes been dubbed "the ratchet effect" (Tomasello et al. 1993). (Tomasello 1999:5)

This process is depicted in Fig. 2.1. In a given environment, individual *A* invents an artefact to perform a desired function. The knowledge of how to make and use the artefact is then passed on from individual *A* to individual *B*. Individual *B* modifies the artefact, maybe to make it perform its function in a novel environment, before it passes its knowledge on to individual *C*. Individual *C* uses the artefact but does not modify it. The artefact is thus passed on to individual *D* in the state in which it had been inherited from individual *B*. Individual *D* adds two modifications to the artefact, after which the knowledge is transmitted further. Each individual thus builds on the achievements of previous generations, and the point from which individual *D* starts might be one that it could not have reached by itself.

There is a plethora of examples for this general process of cultural evolution. Tomasello (1999:37, citation in original) mentions the evolution of hammer-like tools:

> Some cultural traditions accumulate the modifications made by different individuals over time so that they become more complex, and a wider range of adaptive functions are encompassed—what may be called cumulative evolution or the "ratchet effect." For example, the way human beings have used objects as hammers has evolved significantly over human history. This is evidenced in the artefactual record by various hammer-like tools that gradually widened their functional sphere as they were modified again and again to meet novel exigencies, going from simple stones, to composite tools composed of a stone tied to a stick, to various types of modern metal hammers and even mechanical hammers (some with nail-removing functions as well; Basalla 1988).

Similar accounts could be given not just for technological achievements but also for social practices.

The question now arises why humans have cumulative culture and other animals do not. Tomasello (1999:37), referring to Boesch and Tomasello (1998), argues that "[t]here do not seem to be any behaviors of other animal species, including chimpanzees, that show cumulative cultural evolution." Boyd and Richerson (1996:1) make a slightly weaker claim by saying that "cumulative cultural evolution resulting in behaviors that no individual could invent on their own is limited to humans, song birds, and perhaps chimpanzees." At any rate, no animal species, even if it does exhibit traces of it, has evolved cumulative culture to a degree anywhere close to that found in human civilisations.

The answer to the question about the uniquely human trait of cumulative culture has to be sought in the fact that the ratchet effect can only take place if the employed form of cultural transmission is faithful enough to prevent backward slippage. A newly invented artefact has to be preserved until further modifications can be made. This means (i) that it has to *persist* within an individual until it is passed on to another individual, and (ii) that it has to be *transmitted faithfully* to that other individual. Weak forms of social transmission, like the ones discussed above, do not provide such faithful transmission. In these forms of culture, the ratchet constantly slips backward as illustrated in Fig. 2.2. Individual *A* innovates an artefact. By means of weak forms of social transmission such as local enhancement, the discovery made by individual *A* leads to individual *B*

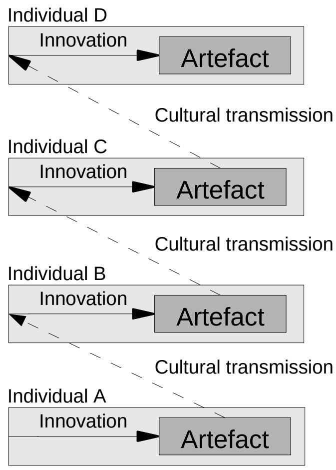


Figure 2.2: The slippage effect in non-cumulative culture. Individual *A* innovates an artefact. By means of weak forms of social transmission such as local enhancement, the discovery made by individual *A* leads to individual *B* making that same or a similar discovery for itself, which in turn triggers individual learning in individuals *C* and *D* respectively. This procedure is opposed to the ratchet effect, which is at work in cumulative culture (Fig. 2.1).

making that same or a similar discovery for itself, which in turn triggers individual learning in individuals *C* and *D* respectively. Tomasello (1999:5, referring to Kummer and Goodall 1985) emphasises the importance of faithful transmission to cumulative cultural evolution:

> [F]or many animal species it is not the creative component, but rather the stabilizing ratchet component, that is the difficult feat. Thus, many nonhuman primate individuals regularly produce intelligent behavioral innovations and novelties, but then their groupmates do not engage in the kinds of social learning that would enable, over time, the cultural ratchet to do its work.

Two reasons have been given for why species other than humans do not succeed at maintaining the cultural ratchet. Tomasello et al. (1993) and Tomasello (1999) provide a qualitative explanation: only humans possess strong forms of social learning, namely imitative learning—Boyd and Richerson (1996) speak of "observational learning"—and active instruction. Boyd and Richerson (1996), however, put more emphasis on the quantitative difference: even though some animals other than humans might in principle be capable of employing strong forms of social learning, they do not engage in it frequently enough in order for the ratchet to be prevented from slipping backward. The result however is the same whether a quantitative or a qualitative explanation is adopted: because non-human animal species do not, or not frequently enough, engage in strong forms of social learning, cultural traits cannot accumulate and cultural evolution like it is known in humans cannot come about. Consequently, Boyd and Richerson (1996:9) venture to conclude that "observational learning seems to require special psychological mechanisms (Bandura 1986) ... shaped by natural selection because culture is beneficial." Tomasello (1999, 2003a) identifies the capacity to understand conspecifics as intentional agents, that is, a theory of mind in the broad sense, as the psychological precondition for imitative or observational learning and points out the inseparability of cultural artefacts from the function they are supposed to fulfil in an environment:

> [T]he understanding of others as intentional beings like the self is crucial in human cultural learning because cultural artifacts and social practices—exemplified prototypically by the use of tools and linguistic symbols—invariably point beyond themselves to other outside entities: tools point to the problems they are designed to solve and linguistic symbols point to the communicative situations they are designed to represent. (Tomasello 1999:6)

The crucial evolutionary step to obtaining cumulative culture is, in this view, the moment where our ancestors began to interpret each other as intentional beings. It is this capacity that enabled them to transmit knowledge with a degree of fidelity which made the accumulation of innovations possible.

2.1.3 Components of cultural evolution

The introduced ratchet-model of cultural evolution comprises three main components: (i) *artefacts*, (ii) the process of *innovation* (or individual learning), and (iii) the process of cultural *transmission* (or social learning). Some further examination of each of them will be necessary if the components are to be identified in the domain of language later on. The remainder of this section is thus dedicated to a more thorough discussion of particular aspects of the introduced model of cultural evolution.

2.1.3.1 Artefacts: function and usage

In the context of cultural evolution, the term "artefact" does not refer to individual man-made objects—this is the sense in which the term is predominantly used in archaeology—but rather to any skill acquired through some form of individual or social learning, that is, through invention or cultural transmission. The most frequently mentioned, quasi-prototypical artefact, the stone axe, can serve as an example here: the product of cultural evolution is not the physical object as such but rather the knowledge of *making* and *using* stone axes. The artefact, in this sense, is thus the learnt behaviour and the knowledge of its application. What evolves is consequently not the individual stone axes themselves but the technique of producing stone axes as well as the way in which they are used to serve their function. This definition allows us to extend the notion of artefact from mere technological achievements to social practices, and in fact any acquired behaviour, including forms of communication and, as I will discuss below, language.

To understand an artefact means to understand its function (the purpose which it is supposed to fulfil in a given environment) as well as its usage (the way in which it is used in a given environment to fulfil this purpose). The *function* of an artefact is defined as the change in the environment which the use of that artefact is supposed to bring about. The function of the macaque's potato-washing is to have the sand removed from them. The function of the young chimpanzee's arm-lifting is to get to ride its mother, and the function of turning logs is to reach the insects hidden underneath. Sperber (1996:99) describes the function of linguistic utterances involved in communication as "to ensure a similarity of content between one of their mental causes in the communicator and one of their mental effects in the audience." (I will provide a detailed discussion of this last, linguistic example in chapter 3.) The *usage* of an artefact is defined by the way in which its user applies it to interact with the environment in order to achieve a goal (the artefact's function).

While function and usage are two aspects that contribute to the form of an artefact in a very obvious way, the constraints imposed by the *environment* must not be underestimated. This can be illustrated by two cases of chimpanzees using wooden sticks to get to termites in their mounds. In eastern Africa, they use small sticks to fish for the termites in the openings of the mounds, whereas in western Africa, chimpanzees use large sticks to destroy the mound. This variation in the use of wooden sticks to reach the termites is not the result of culture but is simply due to environmental conditions: termite mounds in western Africa are of softer consistency than their counterparts in the east of the continent because of the larger amount of rain they are exposed to (Tomasello 1999:28f.). Understanding an artefact thus means to be clear about who is going to use it how and under what circumstances to achieve what: one has to consider *function and usage*, the latter denoting how the user achieves the goal set by the function in a given environment.

2.1.3.2 Innovation

In this thesis, I take the term "innovation" to denote any process which initiates a *cultural change*, that is, the alteration of an artefact. This includes the original invention or discovery of the artefact as well as its later modification. In this sense, innovation is a synonym of the term "individual learning," as it has been used at some places above. Where the distinction is relevant, "innovation" will denote the process and not its product; the latter being referred to as a "(novel) artefact." A thorough conceptual study of innovation from an anthropological perspective has been presented in Barnett (1953).

Innovation is usually characterised as the process of solving a novel task or problem imposed on an individual by its environment. Barnett (1953:97–180) assesses the range of possible incentives for innovation. Kummer and Goodall (1985) discuss the conditions of innovative behaviour in primates. They describe how the environment can occasion innovations, either because it has changed suddenly or because its stability results in an excess of leisure and energy. The presented observations suggest that familiarity with the components of a situation is a prerequisite for innovations of any kind, be they technical or social (Kummer and Goodall 1985:26). This complies with a statement made by Kristiansen (2005:153), who, from an archaeological perspective, argues that one precondition for an invention—which he views as an accumulation of innovations—to take place is the availability of most of its components. Another such precondition he points out is the development of new needs. This observation, in turn, is reflected in the way Tomasello (1999:41) describes the process of how an individual innovates by modifying an artefact on the basis of its original function so that it can meet the current problem situation:

> An individual confronts an artifact or cultural practice that she has inherited from others, along with a novel situation for which the artifact does not seem fully suited. She then assesses the way the artifact

is intended to work (the intentionality of the inventor), relates this to the current situation, and then makes a modification to the artifact.

Innovations can thus be seen as adaptations to a changing environment.

But not just large-scale environmental changes elicit innovations: in principle, each instance of use of an artefact constitutes an elementary innovation. Each time an artefact is used, it is used in a slightly altered situation. Novel conditions do not just arise as the result of large-scale environmental changes but occur in principle with every new context of use which is minimally different from the previous ones. While somebody knows, in general, how to drive a nail into a wall, every time they do so, they are innovating on a very low, elementary level. They might not be using the same nail, and definitely not the same spot on the wall, probably not even the same wall. The course of the hammer blows might necessitate different correction measures, and finally the person's physical condition might not be exactly the same that it was on previous days. Heraclitus allegedly summarised this by stating that no man can ever cross the same river twice: the river will not be the same anymore, and neither will the man. That each application of an artefact to a novel situation—albeit that the "novelty" of this situation appears to be minute—yields an elementary low-level innovation can also be concluded from the description that Barnett (1953:181) provides of how innovations happen: "[w]hen innovation takes place, there is an intimate linkage or fusion of two or more elements that have not been previously joined in just this fashion, so that the result is a qualitatively distinct whole." This linkage, of course, can happen in two ways: an existing artefact is used in a novel context, or a novel artefact is designed to fulfil a certain function in a given environment.

In summary, innovation is a process that is closely related to the function of an artefact because it presupposes the intention to solve of a task for which no adequate artefact is available yet in a specific environment. However, the degree to which innovation is purpose-driven and conscious has been debated at times. Henrich et al. (in press), for instance, criticise Pinker (1997:209) for overestimating the importance of conscious innovation, and add that innovation can often be the result not of the deliberate consideration of a problem but of accidental discoveries. Innovation does not necessarily have to be brought about by somebody racking their brain but can be due to much less conscious processes. On the other hand, even in cases where accidents and happenstance lead to a discovery, the process of cultural change is still driven by a certain degree of intentionality. Only if I plan to solve a problem do I apply an accidentally discovered fact *A* to some pre-existing knowledge *B* to form a novel artefact *C*. A different issue is touched on by those who assume that in cultural evolution, innovation is basically a by-product of transmission, to which I will turn now.

2.1.3.3 Transmission

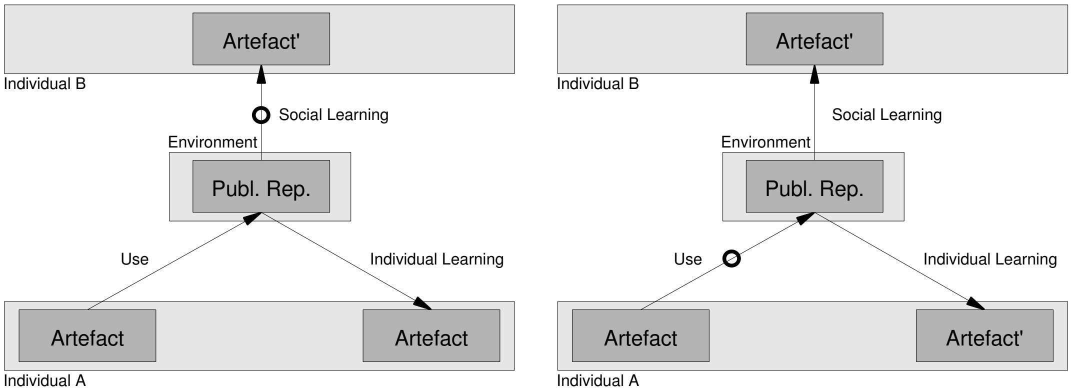
The important question about the process of transmission is how faithful it is. Tomasello (1999:5, and later almost identically in 2002:331f.) emphasises that faithful social transmission is vital to cumulative cultural evolution "so that the newly invented artifact or practice preserves its new and improved form at least somewhat faithfully until a further modification or improvement comes along." Boyd and Richerson (1996) agree with this general point—with the already stated difference in how they explain the absence of cumulative culture in most non-human species: while Tomasello assumes that they do not engage in forms of social learning capable of faithful transmission, Boyd and Richerson argue that even though they might do it, they do not do it frequently enough. This slightly more differentiated view does not, of course, contradict Tomasello's more abstract statement; in fact, he explicitly acknowledges that the situation might be more as described by Boyd and Richerson: "The argument is thus that there is a quantitative difference in social learning skills that leads to a qualitative difference in the historical trajectory of the resulting cultural traditions" (Tomasello 1999:40). By speaking of cultural transmission preserving an artefact "at least somewhat faithfully," Tomasello (1999:33) evidently acknowledges that there might be minor variation during transmission. Nevertheless, the central point of the ratchet model remains the fact that in general, transmission must be faithful for cumulative culture to evolve:

> The metaphor of the ratchet in this context is meant to capture the fact that imitative learning (with or without active instruction) enables the kind of faithful transmission that is necessary to hold the novel variant in place in the group so to provide a platform for further innovations. (Tomasello 1999:39)

Others, however, assign far greater significance to the fact that social transmission might be only "somewhat" faithful. Sperber (1996) argues that the main difference between cultural inheritance and genetic inheritance is that cultural transmission is intrinsically transformational. He points out that cultural information can only be transmitted from one individual to another via the mediation of a "public representation," that is, that an artefact can only be acquired by observing the physical realisation or behaviour to which it gives rise. The underlying artefact cannot directly be inspected but must be inferred from its observable public representation. Such inference leads, in Sperber's view, almost necessarily to a transformation of the transmitted information. This is what a number of philosophers identify as the main difference between genetic and cultural inheritance: genetic inheritance can be described as a copy-process while cultural inheritance relies on inferential transformation (see e.g. also Boyer 1999; Atran 2001) or so-called "reverse engineering" (Kirby et al. 2004). In this sense, faithful transmission (or the replication of mental representations, as Sperber calls it), "when it truly occurs, is best seen as a limiting case of zero transformation" (Sperber 1996:101).

In conclusion, I identify two sub-models of cultural evolution, each of which has a different locus of innovation. They are represented in Fig. 2.3. An artefact becomes modified either when it is inferred from the observed use of it by another individual, as represented in Fig. 2.3(a), or the user modifies the artefact when trying to apply it to a novel situation or environment, as illustrated in Fig. 2.3(b). We have to consider two possible loci of innovation: use and transmission. In the first case, the observer tries to imitate the original artefact but is not faithful in doing so because he has to infer his knowledge about the artefact from its use in a particular situation. Then, innovation comes about because transmission is not entirely faithful. In the second case, the user modifies the artefact to adapt it to a novel situation, and the observer acquires the already modified artefact. In this latter case, the innovation thus spreads to both the observer (through social learning) and the user (through individual learning), while in the former case, innovation is a matter of the observer only.

If they are viewed separately, the two sub-models assign different roles to the process of cultural transmission. If the fidelity of transmission is emphasised and innovation primarily comes about through use, then the role of transmission is mainly to ensure the continuity of innovation beyond the individuals' lifespans. A cumulation of innovations already happens within an individual's lifetime but it is faithful transmission that makes the cultural ratchet possible. On the other side, if transmission is the prime locus of innovation, it becomes itself the crucial motor of cultural evolution. Either type of cultural evolution



(a) Innovation during transmission (b) Innovation during use

Figure 2.3: Two loci of innovation. The locus of innovation is marked with a circle. The horizontal axis symbolises individual learning; the vertical axis social learning. (a) Individual *A* uses an artefact in a given environment. Individual *B* observes this behaviour and tries to imitate it. However, the artefact acquired by observation of the behaviour's "public representation" has undergone some transformation during cultural transmission. The artefact that individual *B* acquires is different from the one that individual *A* used. The artefact of individual *A* is not modified after the usage event: individual learning does not bring about a change in *A*. (b) Individual *A* applies an artefact to a given environment. Because the environment differs from the one in which the artefact was acquired, the artefact now needs to be modified to meet the novel situation. Individual *B* observes the behaviour, and faithfully acquires the modified artefact. At the same time, the novel usage of the artefact also results in a modification in individual *A* through individual learning.

would still be less time-consuming than biological evolution, for three reasons: (i) individuals learn from each other not just once per generation (even though ontogenetic development is certainly the most intensive phase of learning), (ii) individuals do not just learn from one or two other individuals (the parents), and (iii) the innovations themselves need not be random but can, as has been pointed out above, be relatively purpose-oriented.

The two identified sub-models of cultural evolution are, of course, not mutually exclusive. Cultural innovations can occur as the result of adapting an artefact so that it fulfils its function in a novel environment as well as the product of inferential transformation during the acquisition of it by another individual. Given the fact that cumulative culture relies so substantially on faithful transmission occurring frequently enough, it could be argued that the aspect of non-faithful transmission can be neglected if one assumes a certain level of abstraction, but it would appear to be peculiar if the case was the other way round and the aspect of innovation for use were not considered. Yet an overwhelming proportion of existing studies in the cultural evolution of *language* do just that: they focus on alterations brought about by inferential transformation. Later on in this chapter (2.3.3), I will argue that the fact that they focus on this one, seemingly less important aspect of cultural evolution has something to do with the assumptions they make about the nature of grammar and about the locus of linguistic innovation.

The points to be remembered at this stage are these. Humans have cumulative culture because they engage in forms of social learning that allow a faithful transmission of cultural information. Cultural evolution accumulates modifications of artefacts (knowledge, skills, or social practices) which are used in a certain way within a given environment to perform a specific function, that is, to bring about some specific change of that environment. Because every situation of use is minutely different from previous situations, every use of an artefact also entails slight modifications to the artefact itself. To build a model of linguistic cultural evolution, one has to identify these factors in the domain of language. In particular, it has to be determined what the linguistic artefact, its function and its usage is, and where the locus of linguistic innovation lies. The remainder of this chapter is dedicated to these questions: I will map the introduced general model of cultural evolution onto exisiting approaches to the cultural evolution of language.

2.2 Identifying the linguistic artefact

As strange as it seems, it is not common at all for a study of the evolution of language to begin with the question about the very nature of language. Every now and then, a more linguistically-minded member of the language evolution community complains about this fact (e.g. Bickerton 2003; Newmeyer 2003). The assumptions about language itself are often among the most hastily and cursorily made assumptions within a model of language evolution. And more often than not, it is just assumed that linguistics—which is probably viewed in such cases as a unitary field that can be dealt with as a black box—has decided for sure and unanimously what language is and how it is organised. This oddity may be due to two reasons, namely (i) a general reluctance of linguists to enter into speculation about the origins of language (for the historical motivations of this phenomenon see Newmeyer 2003:59–64), and (ii) among those linguists who have ventured to engage in the study of language evolution, the predominance of one particular linguistic school of thought, namely that of generative linguistics.

To make clear one's assumptions about the nature of language—that is, the linguistic artefact, if we want to speak in terms of cultural evolution—is pivotal because a study of the evolution of some entity can only make sense if one knows what that entity actually is. In her essay on what it means to study the transition to language, Wray (2002:2, emphases in the original) puts it like this: "answering the question of *how* and *why* a change occurred depends entirely on your position regarding *what* changed into *what*." Thus, what one has to be clear about, she claims, if one aspires to account for an evolutionary transition is one's assumptions about the "state before" and "the state after." The assumptions one makes with respect to these two states necessarily define and constrain the solution one comes up with. Disagreements between competing theories of language evolution are often not so much due to differing views of what processes applied but rather founded in the assumptions that are explicitly or implicitly made about what language actually is. A theory of language evolution thus has to specify (i) what the evolutionary state without language is (this will be discussed in chapter 3) and (ii) what language, as the evolved entity one intends to explain, actually is. In this section, I will address the latter question by first discussing the general nature of language and then introducing the views held by the two predominant contemporary schools of linguistic thought: generative linguistics and cognitive-functional linguistics.

2.2.1 Linguistic competence

To ask about the linguistic artefact is to ask about a cognitive entity. We have seen in the general discussion of cultural evolution above that artefacts must be viewed as the information or knowledge of how to achieve a specific change in a certain environment. They have been described as skills or practices, and I have illustrated that it is not, for example, the stone axe itself but rather the individual's knowledge of how to make and use a stone axe that has to be considered as the cultural artefact. Artefacts thus reside in what can be called the cognitive environment of an individual, and they have been described, at an abstract level, as mental representations (see e.g. Sperber 1996:32ff.).

Contemporary linguistic theory is in harmony with this claim.[3] Taylor (2002:5) summarises the basic point on which most contemporary branches of linguistics seem to agree as follows:

[3] In fact, it must be assumed that the anthropological characterisation of artefacts has at least partly been influenced by what has been described as "the cognitive turn" in linguistics.

> I think it is fair to say that most linguists, nowadays, would at least pay lip-service to the idea that language knowledge resides in the mind, and that what linguists are trying to do, as linguists, is to describe what is in the mind that enables people to create and understand linguistic expressions.

This, of course, has not always been the case. In fact, Chomsky (1959), which is usually seen as having initiated the cognitive turn in linguistics, is a reaction against the earlier non-cognitive, behaviourist and structuralist paradigm.[4] This earlier approach to linguistics, which had been shaped significantly by Bloomfield (1933), conceived linguistics as concerned with describing the observed behaviour of whole speech communities without any reference to what must be going on in the minds of the individual language users. Chomsky (1965) shifted the focus of interest towards the psychology of the individual speaker. While Chomsky (1957) had introduced generative grammars as a formalism to describe all and only the well-formed sentences of a language, he now claimed that generative grammars had to be seen as a theory of the linguistic competence of an idealised speaker of that language.[5] Consequently, the aim of linguistic theories, whether they comply with Chomsky's approach or not, has become to describe *linguistic competence*, that is, as stated in the above quote, "to describe what is in the mind that enables people to create and understand linguistic expressions."

Modern linguistic theories use several terms related to linguistic competence in an often not clearly distinct, loosely interchangeable way. Chomsky (1986b, 1991) introduces the term "Internalised (I-) language" as standing for the same thing as competence, namely "the speaker-hearer's knowledge of his language" (Chomsky 1965:4). This notion of language as the internalised knowledge of an individual speaker has to be contrasted with three other concepts. The first is, as already mentioned, the idea of language as an abstraction over the linguistic data observed in a community of speakers, "a language" in its everyday sense, which Saussure (1916) calls "langue" and to which Chomsky (1986b, 1991) refers as "Externalised (E-) language." E-language research is characteristic for purely descriptive, behaviourist approaches to language. The second concept is language as the way in which a speaker uses linguistic expressions in concrete

[4]Unlike Chomsky (1965), Chomsky (1957)—published two years before the criticism of Skinner in Chomsky (1959)—does not have a distinctively cognitive character yet but is more clearly based on its structuralist precursors.

[5]It has to be noted that the fact that Chomsky assumes an *ideal* speaker of a particular language in part still reflects the notion of linguistics as the study of a linguistic system as the abstraction over the behaviour of a speech community.

speech situations, sometimes including the mental processes involved in producing and understanding actual utterances. This notion of language is what Saussure refers to as "parole" and what Chomsky (1965) calls "performance" (as opposed to "competence"). Occasionally, Chomsky uses the term "pragmatic competence" (Hymes 1972 suggests the term "communicative competence" for the same concept) when he speaks of the knowledge that "places language in the institutional setting of its use, relating intentions and purposes to the linguistic means at hand" (Chomsky 1980:225). However, we will see in chapter 3 that the distinction between linguistic and pragmatic competence envisaged by Chomsky is not necessarily a useful one and that the borderline he draws turns out to be rather artificial. Finally, what researchers mean when they speak of the evolution of "language" pertains to a third notion of the term. It can be used to denote the phenomenon of linguistic communication in general (Saussure's "langage") and the cognitive equipment by which it is enabled, the "language faculty." A last term which deserves mentioning is "grammar." We will see in section 2.2.3 below, that depending on the linguistic theory, "grammar" is either used as a synonym for linguistic competence or as one particular component of it. In this context, we also have to be aware of the distinction between the general psychological characteristics of human linguistic competence on the one side ("grammar" as a mass noun) and the specific instantiation of this linguistic competence possessed by a particular individual on the other side ("a grammar"). Finally, it will pay to be aware that, depending on the linguistic convictions of the observer, "grammar" can refer to an actual phenomenon with a psychological reality of its own, or merely describe an emergent epiphenomenon in the eye of the observer:

> There is a constant danger of silently moving from [the] understanding of "grammar" as an independent, discrete, mentally represented, innate entity to a more fuzzy conception of grammar as a mass noun, a cover term for a wide and heterogeneous spectrum of regularities in speech as noticed by the linguist. (Hopper 1991:45)

2.2.2 Associating form and meaning

So, what does it mean to know language? What constitutes an individual's linguistic competence? I have pointed out in the general discussion of cultural evolution above that to understand an artefact means to understand its function and usage, that is, *how* it is applied (its usage) to achieve *what* change in the environment (its function). A cursory answer—which will be further specified in chapter 3—to the question of what grammar is and does, on which most contemporary linguistic theories agree, is the following: language is first and foremost an association of (phonological) form and meaning. Chomsky (1972:17), the founder of generative linguistics, states that "each language can be regarded as a particular relationship between sounds and meaning," and Langacker (1987:11), one of the pioneers of cognitive linguistics, asserts that "language is symbolic in nature. It makes available to the speaker [...] an open-ended set of linguistic signs or expressions, each of which associates a semantic representation of some kind with a phonological representation." When I know a language, then I know what particular linguistic expressions in that language stand for, and I know how to express particular meanings in the phonological forms provided by that language.[6] This is, however, about as far as the agreement between generative linguistics and cognitive-functional linguistics goes. The question that the two approaches answer quite differently is *how* language achieves the mapping between form and meaning—and ultimately also *why* it maps the two onto each other in the first place, i.e. whether the function of the mapping is communication or not.

2.2.3 Models of linguistic competence

To think that linguistics has determined one sole description of what language is, and therefore to treat grammar as a black box in one's theory of language evolution, is making simplifying assumptions to a degree where they affect the results. The two rather distinct conceptions of linguistic competence currently dominating the field are generative grammars and construction grammars. The distribution of these two views roughly corresponds to the two linguistic research agendas of generative linguistics and cognitive-functional linguistics. In the following paragraphs, I will, for each of the two interpretations of linguistic competence, (i) provide a brief characterisation, (ii) discuss how it achieves the mapping of form and meaning, and (iii) refer to the main theories that represent it.

[6] NB that the term "phonological" has to be used in its broadest sense, pertaining to duality of patterning and the definition of a phoneme as the smallest unit that distinguishes meaning (without carrying meaning itself), and therefore without necessary reference to sound as such, including sign language just as well as spoken language.

2.2.3.1 *Generative grammars*

The main characteristic of generative grammars is that they divide linguistic competence into two components: a *lexicon* containing symbolic units that combine form and meaning, and a *computational system* of abstract rules that operate on top of the symbolic units (Chomsky 1965:15–18). Chomsky (e.g. 1995b:130) repeatedly characterises this as one of the preliminary assumptions of his theories. "The generative procedure that constitutes language consists of a lexicon and a computational system" (Chomsky 1995b:33, almost identically e.g. in Chomsky 1995a:15). This view of grammar has consequently come to be referred to as the "words and rules" approach (Pinker 1999). The function of the rules is to generate all and only the grammatical sentences of a language. Such rules operate independently on various levels: phonological rules determine the arrangement of sounds, syntactic rules organise words into sequences, and semantic rules constrain possible meanings. Each level has its own autonomous system of rules. At the core of the linguistic competence is thus a computational system—the term "grammar" is often applied to denote this computational system only—whose rules are responsible for the building of the structure that can be observed in language.

The mapping of form and meaning in generative grammars is thus an indirect one. It is mediated by the computational system whose rules *transduce* meaning into form rather than associate them with each other directly. The only direct pairing of forms and meanings is contained in the lexicon, whose entries merely serve as input material for the transduction algorithm. They are the symbols manipulated by the abstract syntactic machinery of grammar. Cook and Newson (1996:43) summarise the generative view on how linguistic competence maps forms and meanings as follows:

> If language could be dealt with as pure sounds and as pure meanings, its description would be comparatively simple. The difficulty of the task is due to the complex and often baffling links between them: how *do* you match sounds with meanings? The answer is the 'computational system' (Chomsky 1993) present in the human mind that relates meanings to sequences of sounds in one direction and sequences of sounds to meanings in the other.

The generative model of linguistic competence is illustrated in Fig. 2.4.

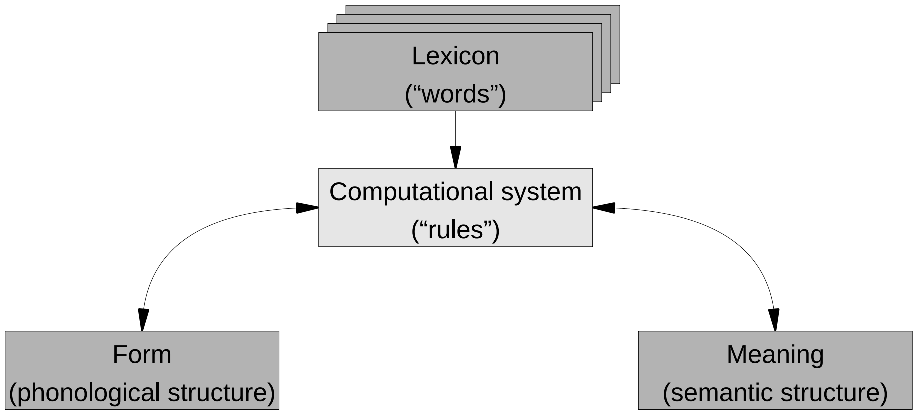


Figure 2.4: The generative model of linguistic competence. Meaning and form are only related to each other in an indirect way. The rules of the computational system, which operate on top of the entries in the lexicon, transduce one into the other. The only place where forms and meanings are associated with each other directly are the words in the lexicon.

The generative view of linguistic competence was introduced in the works of Chomsky (1957, 1965) and received further refinement in the development of his theories such as Transformational Grammar (the Standard Theory, the Extended Standard Theory, the Revised Extended Standard Theory), Principles and Parameters Theory (Chomsky 1981, 1986a,b), which is also known as Government and Binding Theory, and, most recently, the Minimalist Program (Chomsky 1993, 1995b). Theories which adopt the generativist view of linguistic competence also include Lexical-Functional Grammar (Bresnan 1982), Generalised Phrase Structure Grammar (Gazdar et al. 1985), Head-driven Phrase Structure Grammar (Pollard and Sag 1993), and Categorial Grammar (Wood 1993).

2.2.3.2 Construction grammars

Construction grammars view linguistic competence as an *inventory of form-meaning pairings*. This view was first formulated by Langacker (1987:57): "[grammar] can be characterized as a structured inventory of conventional linguistic units." Langacker describes form-meaning pairings as conventional linguistic units consisting solely of a phonological part (their perceptible form), a semantic part (their meaning), and a symbolic relation between the two. With the advent of other theories in the same tradition, the term "construction" has become more common to denote such units (Fillmore et al. 1988; Goldberg 1995).

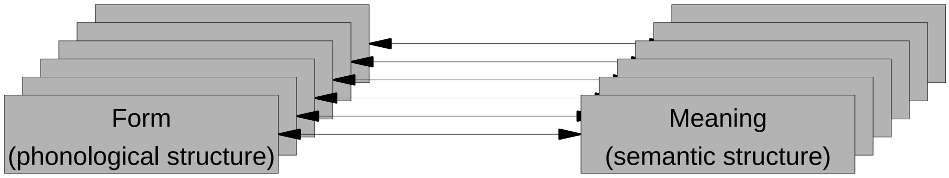


Figure 2.5: The construction-based model of linguistic competence. According to the symbolic thesis, grammar is exhaustively described as an inventory of constructions, that is, as an inventory of individual form-meaning pairings.

Not everybody applies "construction" to *all* form-meaning pairings: some only use it to refer to complex, abstract units, whereas others include atomic, concrete ones such as morphemes and lexical items as well (see e.g. Taylor 2002:567 for a comparison of his more idiosyncratic use of the term with the more common definition employed by Goldberg 1995). In the remainder of this thesis, I will use the term "construction" in accordance with Croft (2001) and Croft and Cruse (2004:256) to denote *any* conventional linguistic unit in the sense of the above definition, that is, for any conventional association of a form and a meaning. The construction-based view of linguistic competence is thus that it can be described exhaustively as an inventory of constructions. This is often referred to as the *symbolic thesis* (e.g. Taylor 2002; Evans and Green 2006). Linguistic competence as an inventory of constructions is represented in Fig. 2.5.

Constructions can be atomic and concrete (specific), or more complex and abstract (schematic). Croft (2001:17) illustrates this with the following examples. Atomic and specific constructions are lexical items such as, for example, [*the*] or [*jumper*].[7] Word classes such as [NOUN] or [VERB] are atomic too, but schematic. Complex constructions can be morphological patterns such as [NOUN-*s*] or [VERB-TENSE], idioms like [*pull*-TENSE NP*'s leg*], or syntactic patterns such as, for example, the English passive construction [NP *be* VERB-*ed by* NP]. Note that complex constructions can contain schematic as well as specific elements. Like this, it

[7]Square brackets signify constructions.

is achieved that the whole linguistic competence of a speaker can be described in terms of constructions. One consequence of this is that construction grammars assume a *lexicon-syntax continuum*: grammatical knowledge is a continuum from concrete/specific to abstract/schematic constructions, and from atomic to complex ones, but it consists entirely of pairings of form and meaning. Complex and schematic constructions are conceived as bearing meaning just like morphemes and words. A seminal empirical study that shows this is Goldberg (1995). In all cases, constructions are immediate associations of form and meaning.

Construction grammar has its origins in the work of Langacker, Fillmore (and associates) and Goldberg. It has developed into a number of individual theories such as Cognitive Grammar (Langacker 1987, 1991; Taylor 2002), Construction Grammar (Fillmore et al. 1988; Kay and Fillmore 1999), the construction grammars developed by Lakoff (1987) and by Goldberg (1995), Radical Construction Grammar (Croft 2001), and, most recently, Embodied Construction Grammar (Bergen et al. 2004; Bergen and Chang 2005) and Fluid Construction Grammar (Steels 2004), two versions of construction grammar developed specifically for computer simulations.

2.2.4 The case for a construction-based view of grammar

The question whether generative grammars or construction grammars provide a more accurate description remains unresolved—a discussion of this would fill a thesis of its own. However, in contrast to what still seems to be a majority of studies into the evolution of language—both nativist and non-nativist, this thesis will be clearly positioned within the construction grammar approach to linguistic competence. In the following, I will briefly sketch the two main reasons for this decision: theory consistence and psychological plausibility.

The generative view of grammar is closely related to the innateness assumption: nativist theories of language usually equate the computational apparatus at the core of the generative grammar with the domain-specific, genetically determined language faculty. This is not to say that a generative view of grammar logically implies the innateness assumption. In fact, computational models of the cultural evolution of language comparatively often resort to generative grammars as an abstract representation of human linguistic competence. However, it seems to me that if one does not want to make the innateness assumption but rather attempts to explain the emergence of language on the basis of domain-general cognitive capacities, it makes sense to also adopt the construction-based view of grammar which has been developed on the basis of the same premise.

The second reason to chose a construction-based over a generative view of grammar is concerned with their respective psychological plausibility. There is a growing body of studies that suggest that the generative view of grammar does not accurately capture the psychological reality of speakers' linguistic competence. One of the most seminal studies in this context is Goldberg (1995), which provides evidence for the fact that, in contrast to what is assumed in the "words and rules" conception of grammar, grammatical constructions carry meaning of their own, independent of the meaning of the words they contain. The view of grammar as an abstract symbol-manipulation system operating over symbolic lexical units can thus not be maintained in the face of the data provided by Goldberg. Other studies demonstrate how generative grammars cannot adequately capture the linguistic phenomena of idioms (Fillmore et al. 1988; Nunberg et al. 1994, Croft and Cruse 2004:ch. 9, Taylor 2002:ch. 27, Evans and Green 2006:ch. 19), collocations (Barlow 2000) and any kind of figurative language. Langacker (1987:29) identifies what he calls the "rule/list fallacy" as one of the reasons why these problems occur, namely "the [false] assumption, on grounds of simplicity, that particular statements (i.e. lists) must be excised from the grammar of a language if general statements (i.e. rules) that subsume them can be established." Such an assumption would imply that an idiom like *cook somebody's goose* has to be characterised as a semantically unanalysable atom. But simple psycholinguistic experiments show that speakers access both the idiomatic meaning of the whole as well as the literal meaning of its parts, as is exemplified in the following dialogue fragment taken from Taylor (2002:551):

(1) A: Did they cook his goose?
B: He doesn't have a goose to cook.

The consequences such evidence has for what is a psychologically plausible view of linguistic competence are interpreted similarly by various proponents of construction-based views of grammar (Lakoff 1987; Langacker 1987, 1991; Fillmore et al. 1988; Goldberg 1995; Fillmore et al. 1988; Kay and Fillmore 1999; Croft 2001). They all agree that psycholinguistic data suggests that I-language can and must be described exhaustively as an inventory of form-meaning associations. The assumption that linguistic competence can be described as a generative grammar can only be maintained if a substantial amount of empirical data is ignored; as a model of I-language, it thus lacks psychological plausibility.

> [T]he requirement of generativity entails the exclusion from the grammar (and hence omission from serious consideration) of both usage and figurative language, which are pivotal to an understanding of linguistic structure. (Langacker 1987:64)

Langacker's statement reflects a fundamental disagreement between generative and cognitive-functional linguistics on the nature of what counts as linguistic evidence, and ultimately what linguistic enquiry is about. Generative linguists typically argue on the basis of intuitive grammaticality judgements of sentences deprived of any context, whereas functionalists maintain that the basis of the study of language has to be actually occurring utterances within their specific context. In his response to Newmeyer (1991), Hopper criticises the data base used in generative linguistics :

> [Newmeyer] unquestioningly assumes that "functional" linguistics and formal linguistics work with the same data base. This data base is to consist of fictitious sentences drawn from linguistic intuitions. [...] The problem is that sentences obtained by consulting "intuitions" are remote from real utterances; they lack a context, they are not socially real, they are literally useless; there cannot be a functional explanation for a non-functional sentence. (Hopper 1991:46)

In his seminal paper, Lakoff (1991) points out that the view that linguistic competence can be described as a generative grammar is not the result of empirical studies, that in fact, it is no *result* at all, but an *a priori* assumption of the generative enterprise. The differences in the two research programs have far-reaching impacts, Lakoff argues, and he states that cognitive-functional linguistics is built on "the commitment to characterize the general principles governing all aspects of human language" (the "Generalisation Commitment") and the "commitment to make one's account of human language accord with what is generally known about the mind and brain from disciplines other than linguistics" (the "Cognitive Commitment"), whereas generative linguistics is committed "to describe language in terms of the mathematics of symbol manipulation systems" (the

"Chomskyan Commitment," Lakoff 1991:53f.). The identification of these commitments leads him to the conclusion that the two research programs cannot actually be compared because they have fundamentally different aims:

> The Cognitive and Generalisation commitment are just commitments to engage in scientific research, whereas the Chomskyan Commitment is a commitment to a program of speculative philosophy: to see what happens if you decide to study language given the metaphor that a grammar of a human language is a symbol-manipulation system in the technical sense. There are a great many linguists who, like myself, were trained as generative grammarians and then moved on to cognitive linguistics because we found that the Chomskyan Commitment was not consistent with what we saw as the scientific study of language. (Lakoff 1991:62)

Even if we do not link the generative model of grammar and language with the innateness hypothesis, the Generalisation and Cognitive Commitment as expounded by Lakoff above appear to match the aim of a culture-based account of the evolution of language much better than the Chomskyan Commitment which underlies the generative enterprise. It thus seems to be more promising to apply cognitive-functional conceptions of grammar in models of the *cultural* evolution of language. And these converge all on one property: they describe language as a redundant inventory of constructions.[8]

I will now explore how the developed model of general cultural evolution can be mapped onto established models of the cultural evolution of the language.

2.3 Models of linguistic cultural evolution

Models of the cultural evolution of language can be viewed as generally formulated models of how linguistic artefacts come into being (*language evolution* in the narrow sense as the object of study of evolutionary linguistics) and how they are modified (*language change* as the object of study of historical linguistics). A model of the cultural evolution of language will thus have to subsume models of language change but differ from them in two ways. First, it will have to be able to account for qualitatively and quantitatively more substantial leaps than

[8]Some recent studies argue that even as biological entities, that is even if they were innate, construction grammars would be evolutionarily more plausible than generative grammars (e.g. Jackendoff and Pinker 2005; Kinsella in press).

the ones described by historical linguistics. A model of the cultural evolution of language will ultimately have to accommodate what I called the "emergence puzzle" in the introduction to this thesis, namely the problem that the starting point of a study of language evolution is not, as it is the case for a study of language change, language but rather a pre-linguistic, that is, non-linguistic state. Chapter 3 of this thesis will be dedicated to the emergence puzzle. However, the maxim of the uniformity of process (cf. section 1.3.2) would demand that a model of the cultural evolution of language be formulated so that the same processes that lead to language change can also be employed to account for the emergence of the linguistic artefact in the first place. This leads us to the second aspect in which a model of the cultural evolution of language differs from a mere model of language change. This second distinction is one of the degree of abstraction, or rather generality. A model of the cultural evolution of language will strive for an account of the minimal set of *general* cognitive mechanisms *necessary* to account for how linguistic artefacts emerge and change. It will thus have to subsume more specific processes of language change but will have to reach a greater level of abstraction and find motivation for the processes it applies in general cognition. Tomasello (2003b:103) calls for such models of the underlying cognitive processes of individual phenomena of language change when he argues that "[s]ystematic investigation into processes of grammaticalization and syntacticization is still in its infancy":

> Exactly how grammaticalization and syntacticization happen in the concrete interactions of individual human beings and groups of human beings, and how these processes might relate to the other processes of sociogenesis by means of which human social interaction ratchets up the complexity of cultural artefacts, requires more psychologically based linguistic research into processes of linguistic communication and language change.

There is thus a certain overlap between models of language change and models of language evolution, and in the discussion below I will refer to examples of both domains when explaining how existing models of the cultural evolution of language map onto the general, not language-specific framework of cultural evolution which I have introduced above.

Both models of cultural evolution as well as models of language change can provide so-called micro-dynamic or macro-dynamic accounts—or, of course, any hybrid thereof. It is therefore appropriate to have, very briefly, a closer look at this distinction and the terminology and concepts involved. Micro-dynamic change refers to the emergence of an innovation in an individual. Macro-dynamic change occurs when such an innovation spreads in a community (Nerlich and Clarke 1992). Traugott and Dasher (2005:35, referring to Weinreich et al. 1968 and Milroy 1993; but see also Milroy 1992:ch. 6) clarify the involved terminology:

> The distinction between individual "micro-dynamic" and communal "macro-dynamic" change is similar to that between "innovation" (which occurs in the individual and may not spread to others) and "change," which involves spread across speakers, communities, and registers.

This distinction is, of course, not restricted to language: any model of cultural evolution can have its focus on *innovation* (micro-dynamic change) or the *spread of change* (macro-dynamic change). Historical linguistics has seen a move from system- or language-oriented to more speaker-oriented frameworks, that is, from macro-dynamic to micro-dynamic accounts. Rather than studying how a change propagates in a linguistic community, and how the system abstracted over the linguistic behaviour of that community changes, researchers investigate how innovations in individuals' linguistic competence emerge in the first place, that is, they study the origin or so-called *actuation* of language change. In the context of grammaticalisation, Kuteva (2001:121) describes innovation—or locus of change, as she labels it—as "the specific, local context in which an existing linguistic expression acquires a new interpretation."[9] An innovation, of course, might only become "visible" to the observer if it persists in the innovator (i.e. if it is used again) and if it is transmitted to other individuals. "An innovation made by an individual speaker can initiate a change. But it will lead to change only if it is adopted and used by a sufficient number of other speakers and is then acquired and used by new cohorts of speakers" (Andersen 2006:67). But the way an innovation spreads in a community does not tell us anything yet about how the innovation emerged in the first place. Fig. 2.3 above, representing the general model of cultural evolution introduced in this chapter, is thus very much

[9] Note that I use the term "locus" in a slightly more general sense than Kuteva, namely as the locus *within the model* of cultural evolution where innovation happens rather than as the specific usage event in which it occurs.

concerned with just these micro-dynamics and the actual locus and nature of *innovation*, as opposed to *spread of change*. In the remainder of this section, I will discuss how this general model can be interpreted for language, and how some of the most predominant types of models of linguistic cultural evolution can be mapped onto it.

2.3.1 Iterated learning

Contemporary models of linguistic cultural evolution are—in one way or another—based on the fundamental distinction between I-language and E-language introduced above. This is, as Kirby (2002a:175) puts it, the simple fact that language exists in two domains: as linguistic competence in users' minds (I-language), and as utterances in actual speech situations (E-language) (Chomsky 1986b; Hurford 1987; Kirby 1999). This dichotomy is equivalent to the one developed above for artefacts on the one side, and the event of putting them into use (i.e. what Sperber 1996:99 calls "public representations") on the other side. It has to be noted at this point that Hurford and Kirby's use of the term "E-language" is slightly different from the one introduced by Chomsky (1986b). As we have seen in section 2.2.1 above, Chomsky's notion of it is related to Saussure's "langue,", that is to language as the system abstracted over the linguistic behaviour of a community of speakers (e.g. English, Greek, Korean). In contrast, Hurford and Kirby's use of the term is closer to Saussure's "parole" and Chomsky's "performance," the linguistic data produced in actual usage events. In the remainder of this thesis, I will use "E-language" in this latter sense, as it is the one more commonly applied in evolutionary models of language.

The central idea of any linguistic model of cultural evolution is the one that I-language produces E-language, which in turn leads to I-language, and so on. Speakers produce utterances on the basis of their linguistic competence. Such utterances then serve as the basis for the linguistic competence of a language learner. Andersen (1973:767) referred to this fact when he pointed out that grammars are not transmitted immediately from one generation of speakers to the next but via what he calls the "verbal output" of a speaker. Kirby and Hurford (e.g. 2002) have taken this idea up and applied it to form a general framework of the cultural evolution of language: the iterated learning model (ILM), where I-language is "learnt" iteratively from the output (E-language) produced by some other I-language. "For a language, or a pattern within a language, to persist

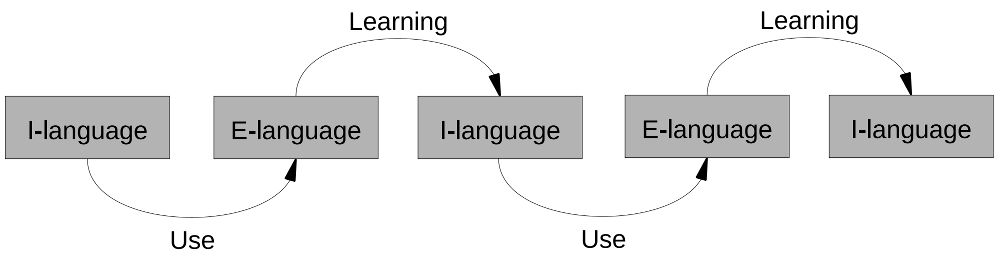


Figure 2.6: Iterated learning. "For a language, or a pattern within a language, to persist from one generation of language users to the next it must be mapped from I-language to E-language (through use) and from E-language back to I-language again (through learning)" (Kirby and Hurford 2002:123).

from one generation of language users to the next it must be mapped from I-language to E-language (through use) and from E-language back to I-language again (through learning)" (Kirby and Hurford 2002:123). This central idea of linguistic models of cultural evolution is represented in Fig. 2.6.

Evidently, the iterated learning model of the cultural evolution of language can easily be mapped onto the general model of cultural evolution introduced earlier in this chapter: I-language is the linguistic equivalent of the artefact and E-language pertains to the process of putting this artefact into use. Thus, if we adapt Fig. 2.3 from above to language, we end up with a situation as represented in Fig. 2.7. First, individual *A* produces E-language on the basis of her I-language. But then, the comparison with our general model of cultural evolution introduces a further differentiation into the general idea of iterated learning: there are two ways in which E-language can be mapped "back" onto I-language, that is, there are two places where "learning" can happen. Either a second individual constructs his own I-language on the basis of the E-language produced by the speaker, or the first individual herself alters her I-language as a result of the usage event. The former is social learning, the latter is individual learning.

At the end of section 2.1.3.3, I concluded that the introduced model of general cultural evolution leaves room for two loci of innovation: use and cultural transmission. Consequently, the model of linguistic cultural evolution shown in

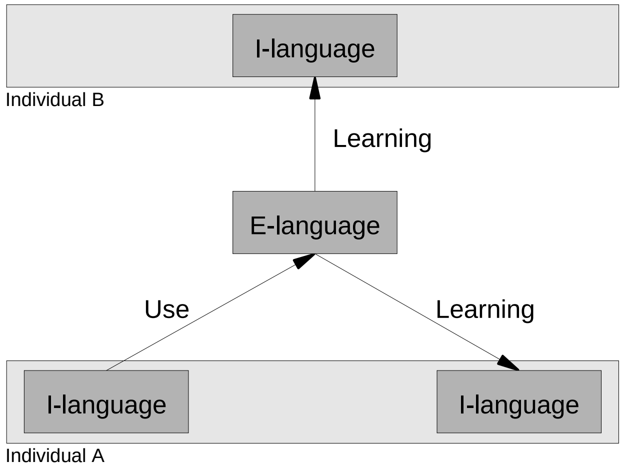


Figure 2.7: Two types of "learning." If the idea of iterated learning, as represented in Fig. 2.6, is mapped onto the general model of cultural evolution described in Fig. 2.3, "learning" must be considered in a more differentiated manner: it is either the mapping of the E-language produced by one individual onto the I-language of another individual, or it is the influence E-language has on the I-language that produced it, that is, on the I-language of the same individual. The former is social learning, the latter individual learning.

Fig. 2.7 also makes it possible for innovations to occur in different positions. I suggest that existing models of language change and evolution can be grouped into three classes, depending on which locus and cause of innovation they focus on.

2.3.2 *Three types of models*

The three instantiations of the iterated learning model that I am going to identify differ from each other in what locus and cause of innovation they emphasise, and what role they assign to communication in this process. It is important to note that these three sub-models of linguistic cultural evolution are not mutually exclusive: it is easily conceivable that the processes described by them exist alongside each other and have all made contributions to the evolution of language. Consequently, some of the actual models that have been proposed contain elements from more than one of these classes of models. The three instantiations of the iterated learning model that I am going to introduce (*acquisition-driven models, comprehension-driven models*, and *production-driven models*) thus have to be seen as idealised "prototypes" that shed light on different possible mechanisms of linguistic cultural evolution rather than as mutually exclusive theories. Nevertheless, identifying their characteristics, in particular where they locate innovation and what causes they assume for it, will help to clarify the issues at stake. For each type, I will provide (i) a brief description, (ii) references to related work in historical linguistics (in particular theories of semantic change and grammaticalisation) on the one side and in evolutionary linguistics on the other side, as well as (iii) a summary of its most defining characteristics.

2.3.2.1 Acquisition-driven models

The notion which is at the core of acquisition-driven models of the cultural evolution of language is the one that E-language only partially specifies I-language. Grammars license an infinite set of sentences but have to be learnt on the basis of a finite set of observed sentences. This finite set may not contain sufficient information for the grammar which produced them to be induced correctly on the basis of the samples alone. This is the idea of *the poverty of the stimulus*, and has more recently been referred to as the *learning bottleneck* (e.g. Kirby 2002a). The processes involved in such a model are thus the *generation* of a limited set of sentences by a grammar and the *induction* of a grammar on the basis of these linguistic samples. Innovations are introduced because of the variability of the observed data set (and the variability of the order in which the individual sentences are presented to the learner) and due to the fact that the learner has to induce a grammar on the basis of a data set which only incompletely specifies it (Fig. 2.8). This interpretation of linguistic cultural evolution thus evolves most crucially around the notion of grammar induction (cf. Wolff 1982), and has consequently been labelled "expression/induction model" (e.g. Hurford 2002).[10] The underlying assumption is that children acquiring language induce grammars in the described way.

Historical linguistics. Acquisition-driven models are particularly popular in generative approaches to language change. McMahon (1994:108) states that "[i]n early Generative syntax, all syntactic change was analysed as simplificatory grammar change" and that such explanations were based on the assumption that "children could construct a different, simpler grammar from their parents'." Similarly, Haspelmath (1998:316) explains that "[s]ince Chomskyan linguistics is primarily concerned with the problem of language acquisition [...], it is natural that generative linguistics should have focused on those aspects of language

[10]Note that Hurford (2002) uses his term "expression/induction model" as a synonym for "iterated learning model." However, Hurford's term seems to be less generally applicable as it clearly specifies the process of learning as (grammar) induction. It must thus remain confined to the described acquisition-driven models of linguistic cultural evolution.

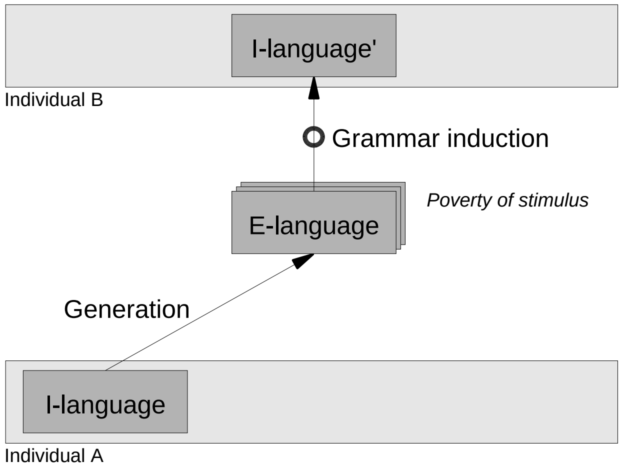


Figure 2.8: The acquisition-driven model of the cultural evolution of language. Individual *A* uses its I-language to generate a set of linguistic data on the basis of which individual *B* infers its own I-language. Because the observed data set only ever partially specifies the I-language which generated it (the poverty of the stimulus) and because individual *B* strives for a concise grammar, cultural transmission introduces innovation (marked with a circle) during grammar induction as the result of imperfect inference.

change that seem to be due to the transmission of grammars to successive generations." Probably the most prominent acquisition-driven model of language change is the one developed by Lightfoot (1979). He argues that a major restructuring of grammars happens in the language acquisition of one generation if grammars have become too complex due to an accumulation of minor changes. In such a situation, children, rather than inducing the fairly complex grammar of their parents, would reconstruct a much simpler grammar on the basis of the observed linguistic samples. Innovation is thus introduced from one generation to the next because children aim at inducing a grammar as simple as possible, which will introduce new regularities where there were exceptions in the consequently more complex grammars of the parents. McMahon (1994:116–37) provides a comprehensive discussion of Lightfoot's proposal, its critics, and subsequent modifications as put forward in Lightfoot (1991). Croft (2000:44ff.) traces the origins of what he calls "the child-based theory of language change" back to 19th-century linguistic theory as described in Jespersen (1922:161–2) and identifies Halle (1962) as its first major instantiation within the generative framework. He also provides a discussion of the problems of this approach, to which we will return in section 2.3.3.1 below.

Evolutionary linguistics. Even though the acquisition-driven model to language change is one developed within the primarily nativist framework of generative linguistics, it has proven to be particularly popular with non-nativist, culture-based approaches to language evolution. Hurford (2002) discusses four such models implemented as computer simulations (Batali 1998, 2002; Kirby 2000, 2002a). Similar approaches can be found in Brighton (2002); Zuidema (2003); Hoefler (2006b). An overview of the work being carried out in this line of research, and of some of the issues involved, is given in Kirby and Hurford (2002). Related to this work is Roberts et al. (2005), who present an acquisition-driven model which, like the models in Brighton (2002) and Hoefler (2006b), makes specific use of the notion of learners striving for grammar simplicity. Grammar simplicity is modelled in the form of the so-called Minimal Description Length principle (Rissanen and Ristad 1994). Niyogi and Berwick (1995, 1996, 1997) develop mathematical studies based on an acquisition-driven model of the cultural evolution of language. Here too, the core notion is that "[t]he problem of language learning [...] is typically formulated as a search by a learning algorithm for a grammar that is close to the one that generates the sentences the learner is exposed to" (Niyogi 2002:231). Another researcher who emphasises the role of acquisition in language evolution is Deacon (1997). He claims that "[t]he structure of language is under intense selection because in its reproduction from generation to generation, it must pass through a narrow bottleneck: children's minds" (Deacon 1997:110), and then states that "[c]hildren selectively hear some structures and ignore others, and so provide a major selection force for language structure that is 'child-friendly' " (Deacon 1997:111).

A similar model has been proposed by Wray (2000), who describes language evolution as a process of segmentation of holistic protolanguage utterances (which stand for whole propositions) along chance co-occurrences of components of form (phonetic segments) with components of meaning: "[t]he result would be the first stages of segmentation: the dividing up of unanalysed material into meaningful subunits, something which has been observed in both first- and second-language acquisition" (Wray 2000:296f.). This is more or less the scenario typically applied in Kirby's computer simulations (see e.g. Kirby 2000, 2002a; Kirby and Hurford 2002). In these studies, compositional syntax is the result of the induction of grammar rules on the basis of matches between segments of form and aspects of meaning, as well as the subsequent merger of grammar rules with the aim to make the grammar more concise (a detailed description of the employed algorithm can be found in the appendix of Kirby 2002a). Note that even though such computer simulations are typically associated with the term "iterated learning model," they represent only one specific, acquisition-driven interpretation of the general iterated learning framework of the cultural evolution of language formulated by Kirby and Hurford (e.g. 2002).[11]

Characteristics. What are thus the characteristics of acquisition-driven models of the cultural evolution of language? Acquisition-driven models are special in (i) where they locate innovation and (ii) the way they assume innovation happens, and in (iii) the role they assign to communication. The locus of innovation in these models is cultural transmission via language acquisition, and the emergence of innovation itself is driven by two concepts: the poverty of the stimulus and grammar simplicity. Not only do children have to infer their I-language from sets of linguistic data which only partially specify the I-language from which they were generated (the poverty of the stimulus) but they also attempt to come up with concise grammars (grammar simplicity). Innovation can thus be the result of generalisations made on the basis of the observed data, or due to generalisations made to compress the inferred grammar after the addition of a new rule.

The type of explanation provided by acquisition-driven models does not refer to the specifics of communication. Rosenbach and Jäger (2008) point out that in a typical acquisition-driven approach such as Kirby (1999), "language usage only very indirectly plays a role, acting as a filter on the input children encounter in first language acquisition, on the basis of which they construct their grammars." They further comment that in doing so, changes in linguistic competence are relegated to first language acquisition (rather than being ascribed to performance). In his overview of four computational models of the acquisition-driven type, Hurford (2002:304) lists "no effect of communication" as one of the defining properties of these models, and elsewhere, Kirby and Hurford (2002:144) argue that their account of the cultural evolution of language "does not make any reference to communication," and that the benefits which emerging structures may have for communication are "merely a fortunate side-effect." While in these models, it is relevant *that* communication takes place because language is transmitted culturally and acquired by new generations on the basis of its externalised

[11]One may want to distinguish between the ILM in the broad sense (ILM-B) as the general framework represented in Fig. 2.6 and the ILM in the narrow sense (ILM-N) as the class of acquisition-driven instantiations discussed here.

form (E-language), the innovations they describe are not based on *how* communication works, but on the selection of observed linguistic data from which I-language (which it only ever partially specifies) is induced. This perspective on the cultural evolution of language resembles, to some degree, the Chomskyan stance that communication does not play an essential role in explaining language (e.g. Chomsky 1980:230).

In summary, we can state that even though acquisition-based models of the cultural evolution of language do not (or at least not necessarily) assume the innateness hypothesis, they have adopted several concepts developed in and closely related to the (nativist) framework of generative linguistics: (i) language acquisition (or, more specifically, grammar induction) as the locus of innovation, (ii) the poverty of the stimulus and a bias for concise grammars as the causes of innovation, and (iii) the fact that communication does not play a role in the provided explanation of linguistic cultural evolution.

2.3.2.2 Comprehension-driven models

Comprehension-driven models of the cultural evolution of language place the act of communication at the centre of their theory. They crucially depend on the notion that meaning is not transferred from the speaker to the hearer directly but that communication always also involves a fair amount of *inference from context*. A hearer infers at least parts of the meaning of a linguistic signal from his cognitive environment, that is, his general knowledge and perception of the world and the conditions of the speech situation. The core idea of comprehension-driven models is that the cognitive environment of the speaker and the hearer will hardly ever be exactly the same. This phenomenon has been described as a "*mismatch* between speaker's and hearer's discourse world knowledge" (Kuteva 2001:169ff., emphasis added). Because of such a mismatch of cognitive environments, a hearer may infer a meaning for a given signal which is slightly different from the meaning intended by the speaker. This may go unnoticed if communication does not fail, and consequently lead to a *misinterpretation* of the signal in the hearer. The signal is mapped onto a new or altered meaning—a process which only becomes evident once the hearer re-uses the same signal in a new way in which it could not have been used by the original speaker. This second type of model of the cultural evolution of language thus evolves around the fact that the comprehension of linguistic signals in successful communication does not necessarily imply that what the hearer understands is exactly what the

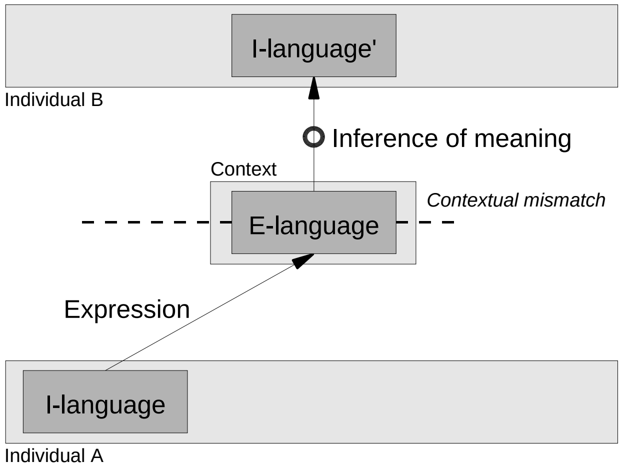


Figure 2.9: The comprehension-driven model of the cultural evolution of language. Within a particular context, individual *A* expresses a linguistic signal to trigger the inference and construction of the meaning to be communicated. However, because the context of individual *B* (the hearer) differs from that of individual *A* (the speaker), the form-meaning mapping inferred by *B* is different from the one that *A* had in mind. Innovation in I-language is thus introduced during the cultural transmission because of imperfect inference due to a contextual mismatch between speaker and hearer.

speaker had in mind. The comprehension-driven model of the cultural evolution of language is represented in Fig. 2.9.

Historical linguistics. At first glance, the comprehension-driven model seems to be closely related to the two processes of language change known as *reinterpretation* and *reanalysis*. Reinterpretation refers to the process by which "the referent of a linguistic form is shifted from one thing to another which occurs in the same context" (Trask 2000:279) or by which the categorial status of a linguistic form is shifted as a result "from its occurrence in ambiguous positions" (Trask 2000:280). Similarly, reanalysis is defined as the process by which a morphological or syntactic structure "comes to be assigned a different structure from the one it formerly had, with no change in its surface form and with little or no change in interpretation" (Trask 2000:274). For a survey of the latter see e.g. Harris and Campbell (1995). The term "reanalysis" has also been applied to refer to acquisition-driven scenarios (e.g. Andersen 2006; Lightfoot sometimes uses the term "radical reanalysis" in his approaches) and there are indeed some parallels between the two models as I will point out in section 2.3.3.1 below.

Trask (2000:274 and 279) provides example (2) to illustrate reinterpretation and (3–6) to illustrate reanalysis. In (2), the meaning of *bead* shifts from denoting 'prayer' to standing for the small balls on a rosary (since prayers are sometimes counted on a rosary) and then simply for any small ball. In (3), the single-morpheme noun *bikini* is reanalysed as containing the element *bi-* 'two' and thus consisting of the constituents *bi-* 'two' and *kini* 'swimming costume'. The fact that such a reanalysis happened would then become clear, once a new creation such as *mono-kini* 'one-piece swimsuit' occurs (4). Similarly, the bracketing in sentence (5) is reanalysed, which becomes visible once sentences such as (6) occur, which would not have been possible without the reanalysis in (5).

(2) bead 'prayer' > 'small ball on the rosary'> 'small ball'
(3) bikini > bi-kini
(4) mono-kini
(5) It will be [easy for us] to do that. > It will be easy [for us to do that].
(6) [For us to do that] will be easy.

However, in this context, it is essential to point to the difference between reanalysis or reinterpretation as the *result* of change, and reanalysis or reinterpretation as a cognitive *mechanism* initiating it. Only if understood in the latter sense can the two concepts be said to represent the comprehension-driven model of the cultural evolution of language; reanalysis or reinterpretation as mere descriptions of the result of change are, as concepts, agnostic as to what sort of innovation they are the result *of*. In order to emphasise this crucial distinction, Kuteva (2001:167f.) issues the following words of warning:

> [E]stablishing that a particular linguistic reinterpretation has taken place does not mean that we have also understood *how* this reinterpretation has taken place. Even when we speak of semantic "bleaching," generalization, metaphor, metonymy, etc. as mechanisms of change, what we are actually talking about is the observable result that is based on *our* comparison between distinct uses, the historically earlier and the historically later. [...] What remains uninvestigated is the unobservable aspect of the psychology of language use, i.e. the psycholinguistic mechanism that triggers the process of semantic reinterpretation in natural discourse.

Within the domain of language change in general, and grammaticalisation theory in particular, Kuteva (2001) offers probably the most clearly formulated comprehension-driven model. She argues that new grammatical functions can result "from *non-shared* discourse world knowledge [...] and from a *mismatch* between the speaker's implicatures/communicative intentions and the hearer's inferences/assumptions" (Kuteva 2001:131) and emphasises the asymmetric relationship between speaker and hearer in an act of communication. In this approach, the focus is thus on knowledge that the hearer only *assumes* he shares with the speaker but that in fact does not constitute part of the two interlocutors' common context. Even though her analysis is mainly concerned with auxiliation, i.e. the grammaticalisation process by which full verbs assume the function of auxiliaries, Kuteva proposes her comprehension-driven model as a general framework for grammaticalisation initiated by mismatches of discourse knowledge between speaker and hearer, and emphasises the central role of such mismatches for everyday communication (Kuteva 2001:167ff.).

Evolutionary linguistics. One of the most comprehension-driven accounts of language evolution is the one proposed by Burling (2000). Burling proclaims the priority of comprehension over production, namely that comprehension is the evolutionarily more basic and powerful mechanism. He bases this hypothesis on the view that "comprehension runs consistently ahead of production" (Burling 2000:27), that is, that humans (and, as he claims, animals) can in any situation understand more than they can express. Furthermore, he emphasises that sometimes, humans (and animals) interpret the behaviour of another individual as a signal even if that individual did not intend to communicate. He infers from this that comprehension came before communication was intended. First signals thus would have evolved simply through individuals interpreting other individuals' behaviour as meaningful. In this account, the first word-like signs were not produced with communicative intention; their communicative value came from the comprehender, not from the producer. In this way, Burling attempts to develop a scenario of the cultural evolution of language parallel to the biological evolution of animal signalling (Maynard Smith and Harper 2003), where behaviour can attain the status of signalling simply through genetic codification of reactions to that behaviour through natural selection. He contrasts this "phylogenetic ritualisation" in biological evolution with the "ontogenetic ritualisation" through individual learning and conventionalisation in cultural evolution (Burling 2000:31). As it is typical for comprehension-driven models of the cultural evolution of language, we can identify a mismatch between the discourse knowledge of the producer, who, in this case, does not intend to communicate, and the comprehender, who assumes that the producer did communicate. Innovation is thus the result of a misinterpretation on the part of the hearer or comprehender.

Another strand of language evolution research which is very explicitly based on a comprehension-driven model can be found in Smith (e.g. 2003a, 2005a,b, 2006a). One of the crucial points Smith makes is that communication can be successful even if the speaker and the hearer do not have exactly the same conceptual structures: a certain mismatch in discourse knowledge will not prevent communicative acts from functioning. Like Kuteva (2001) in her work on grammaticalisation, he too goes a step further and uses computer simulations not only to corroborate this first claim but also to show that evolution of meaning and complexification of language can arise from the dynamics caused by just these differences between the knowledge of the speaker and the knowledge of the hearer. In Smith (2006a), this approach is extended to account for phenomena both in protolanguage theory and in grammaticalisation. In this study, Smith (2006a:310) describes reanalysis as "the unconscious yet inevitable result of the uncertainty involved in the hearer's inferential reconstruction of meaning."

Smith's work has links to two other research programs that deserve mentioning. First, Smith (2005b:380f.) points to the similarities of his computer simulations to those developed by Steels (e.g. 1996b, 1997). These latter models are mainly concerned with the emergent categorisation and the question of symbol grounding in experiments where robots try to establish communication on the basis of their visual perception of the environment. The second line of research to which Smith (e.g. 2005a:322) explicitly refers is Origgi and Sperber's (2000; 2004) account of language evolution. Their argument is also built on the observation that communication does not necessarily fail just because the semantic representation of the two interlocutors are not identical. Origgi and Sperber (2000) apply this fact to situations where, of two communicating individuals, one possesses a more elaborate interpretation of the conveyed message due to some biological mutation. Origgi and Sperber's model thus contains a nativist component: symbols evolve culturally, but grammar is the result of an additional biological step. The transition from the first to the second stage is possible because inferential communication is "failsafe" even where the cognitive representations of two interlocutors exhibit a mismatch. However, it is crucial to note here that Origgi and Sperber's scenario does not depend on the innateness assumption in order to work; the differences in the interpretive capacities of speaker and hearer could simply be empirical, that is, the result of different previous experience.

Characteristics. Like acquisition-driven models, comprehension-driven models too have a particular, characteristic perspective on (i) the locus of innovation,

(ii) the cause of innovation, (iii) the role of communication. Innovation, in comprehension-driven models, is introduced as the result of imperfect inference on the part of the hearer. Like in acquisition-driven models, the locus of innovation is thus in cultural transmission. However, the fact that the recovery of meaning in comprehension is not just a matter of decoding a linguistic signal but also of inferring from context plays a crucial role. Inaccurate inferences leading to novel form-meaning mappings in the hearer occur because there may be a mismatch between the cognitive environment of the hearer and the speaker. The fact that such mismatches occur, but do not necessarily bring down communication completely, is the cause of innovation in comprehension-driven models. The role of the actual communicative act and the impact of its context is thus vital in these approaches.

2.3.2.3 Production-driven models

The notion that communication includes inference from context also plays a central role in production-driven models of the cultural evolution of language. The insight at the core of production-driven models is that language *use* is itself intrinsically innovative. Every event of putting bits of linguistic code into use in a novel situation can be conceived as an innovative act. Bloomfield (1933:433) comments on language use: "[s]ince every practical situation is in reality unprecedented, the apt response of a good speaker may always border on semantic innovation." Innovation, in this view, is the product of the communicative strategies employed by speakers in concrete usage events. Such strategies exploit the fact that language exhibits what I refer to as *pragmatic plasticity*: even though linguistic forms are conventionally associated with meanings (signal meaning), they exhibit flexibility with regard to the meaning they actually come to communicate in a given context (speaker meaning). Croft (2000:99ff.) speaks of the "openendedness and flexibility of meaning in use." In production-driven models, innovation thus enters language not during learning but in use. It is the product of the existing (conventional) meaning of linguistic signal interacting with the context specific to a particular usage event. The process of use can be conceived as the *extension* of existing form-meaning pairings to novel contexts. Since production-driven models move the burden of innovation to use, relative simple forms of learning suffice for these models, such as *memorisation* or *entrenchment* of occurred usages (cf. section 4.1.3). The production-driven model of the cultural evolution of language is represented in Fig. 2.10.

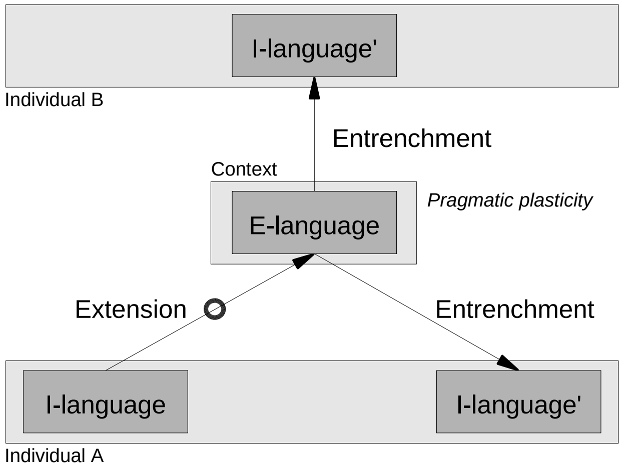


Figure 2.10: The production-driven model of cultural evolution of language. With respect to future usage, constructions exhibit pragmatic plasticity. The process of language use is an extension of existing form-meaning pairings to novel contexts. Since the weight of reasoning and the locus of innovation are in use, learning (in the hearer *and* the speaker) takes the relatively simple form of entrenchment.

Historical linguistics. Studies in grammaticalisation such as, for example, Heine et al. (1991); Hopper and Traugott (2003) have not always been clear about whether they favour comprehension-driven or production-driven explanations. This is mainly due to two facts I have already mentioned above: (i) the lack of a clear distinction between process and result, and (ii) a focus on the description of individual linguistic phenomena rather than on the underlying psychological mechanisms. Hopper and Traugott (2003:71) acknowledge these two facts, yet still take a middle position regarding their model of innovation. They emphasise that grammaticalisation theorists are especially interested in the role of speakers and hearers negotiating meaning in communicative situations (as opposed to the role of language acquisition, as favoured by generative linguistics). This means that they focus on comprehension- and production-driven models, rather than on the acquisition-driven alternative. However, within these boundaries, Hopper and Traugott (2003) do not take sides but rather present a hybrid view that leaves room for comprehension- as well as production-driven explanations:

> On this view, hearers play a major role in change because they process input in ways that may not match the speaker's intention. But speakers also play a major role in enabling change, because in producing

speech they have communication as their goal, and therefore are always in search of ways to guide the hearer in interpretation. (Hopper and Traugott 2003:71)

On the other side, many models of language change developed within the functionalist branch of linguistics seem to be production-driven, like, for example, the seminal *emergent grammar* theory expounded in Hopper (1987), which claims that grammatical structures originate from discourse-pragmatic strategies employed by the speaker.

Other recent studies are more explicit about the issue. Traugott and Dasher (2005) present a seminal study which comes out in favour of a production-driven approach. "The hypothesis is that innovation and change does not occur primarily in the process of perception and acquisition, but rather in the process of strategic choice-making on the part of SP/W [speaker/writer] and interactional negotiation with AD/R [addressee/reader]" (Traugott and Dasher 2005:42). The label Traugott and Dasher use for their approach is *Invited Inference Theory of Semantic Change (IITSC)*. They claim that speakers innovate in communicative situations to bring their message across to the hearer. This view is based on Nerlich and Clarke (1992:127) who relate processes of semantic innovation to the expressive needs which arise in communicative activities. Traugott and Dasher (2005:34) identify metaphorisation (i.e. the novel, ad hoc metaphorical use of lexemes) and metonymisation (they include invited inferencing, subjectification, and intersubjectification in this category) as the innovative mechanisms of language use. They emphasise the speaker-orientedness of their approach by stating that "[i]f SP/Ws innovate, and AD/Rs replicate this innovation, they do so in the role of SP/Ws, i.e. as language producers, not as language perceivers" (Traugott and Dasher 2005:38). In support of their claim that speakers are the prime initiators of change, Traugott and Dasher (2005:19ff., 49) provide evidence for regularity in semantic change towards greater subjectivity or grounding in speakers' attitude and perspective. This can be exemplified with the development of the English phrase *be going to*, which regularly serves as one of the standard examples of grammaticalisation (e.g. Heine et al. 1991; Kuteva 2001; Hopper and Traugott 2003). Subjectification occurs in the shift of the meaning of *be going to* from a merely declarative, objective expression of spatial movement to an expression of modality, that is, an indication of the speaker's attitude and intention.

At the core of Traugott and Dasher's model of innovation is the concept of *invited inferences* (usually called "conversational implicatures" in Gricean pragmatics): speakers innovate by creatively exploiting invited inferences in the particular contexts of speech situations. The term "invited inference" was chosen to stress the speaker-oriented nature of the approach and is defined as follows:

> In the present context it [the term "invited inferences"] is meant to elide the complexities of communication in which the speaker/writer (SP/W) evokes implicatures and invites the addressee/reader (AD/R) to infer them. We prefer this term over, e.g. "context-induced inferences" (Heine et al. 1991), since the latter term suggests a focus on AD/Rs as interpreters and appears to downplay the active role of SP/Ws in rhetorical strategizing, indeed indexing and choreographing the communicative act. (Traugott and Dasher 2005:5)[12]

Traugott and Dasher thus establish a direct connection between the innovative nature of production and the inferential nature of communication. I will explore the underlying psychological mechanisms of this connection between innovative use and inferential communication in detail in chapter 3 of this thesis. In a summary of their model of semantic change, Traugott and Dasher describe equivalents of the processes of extension and entrenchment as depicted in Fig. 2.10 above, that is, the innovative production of E-language and how this process later leads to an innovation in I-language, as follows:

> Drawing on and exploiting, sometimes consciously, sometimes unconsciously, pragmatic meanings, most especially those kinds of implicatures that we have called "invited inferences," SP/Ws may innovate new uses of extant lexemes. If these new uses spread to AD/Rs and are replicated by them in their role as SP/Ws, then semanticization will take place. (Traugott and Dasher 2005:279)[13]

Evolutionary linguistics. Given his general model of cultural evolution as introduced above, Tomasello is the most obvious candidate for a proponent of

[12]Traugott and Dasher (2005:5) clarify, however, that despite the putative implications of their terminology, Heine et al. (1991) also view speakers rather than hearers as the driving-force of innovation.

[13]This particular quote refers to the entrenchment in the hearer, but the entrenchment in the speaker, as also shown in Fig. 2.10, is acknowledged as well, for instance in Traugott and Dasher (2005:35).

the production-driven model within evolutionary linguistics. In his account of language evolution, developed in Tomasello (1999, 2003a) and summarised in Tomasello (2003b), he explains the origin of symbolism as the result of humans understanding their conspecifics as intentional beings such as themselves, and of them trying to manipulate their conspecifics' mental states. It is the mechanisms of such interactions which eventually lead to the conventionalisation of a set of communicative behaviours, one of them being symbolic or linguistic communication. In his approach to the subsequent emergence of grammar, Tomasello takes a position similar to that of Hopper (1987): grammatical constructions (Tomasello subscribes to a construction-based view of grammar as promoted by Langacker 1987, 1991 and Goldberg 1995) emerge from discourse patterns. However, the categorisation of Tomasello's explanation of language evolution as production-driven must remain based on the way he grounds it in the specifics of his ratchet-model of general cultural evolution. This uncertainty about the *linguistic* specifics of his account is due to two peculiarities of his approach. First, it is based on his studies of child language acquisition, and therefore draws its examples from situations different from that of language evolution in the sense that language is already established in the environment. Second, and Tomasello (2003b:103) points this out himself, the processes of grammaticalisation to which he refers to account for the emergence of grammar are still awaiting a more detailed description of the underlying psychological mechanisms—something which I will endeavour to do in chapter 3. These caveats notwithstanding, Tomasello does provide clear evidence for the production-driven nature of his model of linguistic cultural evolution when he describes that the process of innovation is principally the same for material tools (in this case hammers) as well as for linguistic artefacts:

> The first form of sociogenesis is simply the form implied by the ratchet effect as described above for such things as hammers and linguistic symbols. An individual confronts an artifact or cultural practice that she has inherited from others, along with a novel situation for which the artifact does not seem fully suited. She then assesses the way the artifact is intended to work (the intentionality of the inventor), relates this to the current situation, and then makes a modification to the artifact. (Tomasello 1999:41)

Two experimental studies of emergent communication systems deserve brief mentioning in this context too. Galantucci (2005) reports experiments in which subjects are forced to communicate via a graphic channel that prevents the use of conventional symbols and drawing. Garrod et al. (2007) describe how subjects who play repeated games of pictionary develop conventionalised graphical communication systems. Both studies investigate how such systems emerge from a need to communicate, that is, the pressure on the speaker or communicator to bring a meaning across to the hearer or addressee in a situation in which no established code is at hand. In both experiments, speakers innovate on the basis of common ground with the hearers, and the emergent systems reflect properties of the specific environments as well as to the interactive behaviour of the interlocutors. Both experiments also show a gradual shift from very iconic to more arbitrary communication systems. I will come back to this latter point in chapter 3 when I discuss the emergence of a linguistic code from non-linguistic communication.

Characteristics. To analyse the characteristics exhibited by production-driven model of the cultural evolution of language, we can again refer to its views on (i) the locus of innovation, (ii) the cause of innovation, and (iii) the role of communication. As opposed to their acquisition- and comprehension-driven counterparts, production-driven models locate innovation in the process of language use. Ultimately, innovation can occur because every instance of use is intrinsically innovative as it applies an extant linguistic artefact to a novel environment or context. The cause of innovation in production-driven models is thus pragmatic plasticity, that is the fact that, in the specific context of an utterance, linguistic forms take on meaning which goes beyond the meaning that has conventionally come to be associated with them. Language use exhibits "context-dependent deviation" (Kuteva 2001:178); it is an extension of previous usages of a linguistic form to a novel situation.

Production-driven models of the cultural evolution of language can content themselves with relatively simple notions of learning because innovation is introduced elsewhere (in use). Of the three presented types of models, they are thus the ones that come closest to the ratchet-model of general cultural evolution introduced before since they also maintain the fidelity of social transmission. The process of learning in production-driven models can be reduced to simple memorisation of observed usages, or their gradual entrenchment if they occur more often. Since innovation already happens in use, it is later realised in the hearer (through social learning) as well as in the speaker (through individual learning).

Production-driven models are functional: communication plays a pivotal role. Innovation is the product of a communicative intention. We have to be careful, however, to distinguish between two uses of "function," namely (i) the discourse-pragmatic function in a context-dependent communicative interaction from (ii) what might be called "systemic function," that is, the "need" to fill a gap in a grammatical system. Innovation in production-driven models of the cultural evolution of language is *not* the invention of a code or the improvement of a communication system as a whole but the innovative use of an extant bit of code to fulfil a *momentary communicative* need. Comrie and Kuteva (2005) make a case against the system-related type of "functional need" explanations. Likewise, Hopper and Traugott (2003:74) comment on the misconception of the goal-directedness of grammaticalisation through communicative strategies employed in speaker-hearer interactions as being related to grammatical systems:

> Part of the problem with the concept of goal-directedness is that it is often discussed in terms of "need" or set goals, in other words, in "teleological" terms. Clearly, "communicative need" [in the system-related sense] is not a plausible motivation in most cases of grammaticalization, since not all languages express the same grammatical functions, and even less do they express the same grammatical functions in the same way. [...] The position we take is that users may be consciously or unconsciously goal-oriented (see Keller 1994 for a detailed account of goal-orientedness in language change that is not teleological in the sense mentioned above). In speaking of communicative strategies and problem solving in the course of speaker-hearer interaction, we refer not to filling gaps, but rather to strategies used by speakers and hearers in producing and understanding the flow of speech as it is created.

A last characteristic to be mentioned here is the relation between what is the default and what an exception in language use. The production-driven view of language use assumes that "context-dependent deviations" are the normal case of language use, that is, that any instance of use is intrinsically an extension of the existing I-language. This is the inversion of the standard generative view of language use, where deviations from the "norm" are the exception rather than the default case. The conception of the default character of innovative use is shared by Sperber and Wilson (1995). Within the framework of Relevance Theory, they

	acquisition-driven models	**comprehension-driven models**	**production-driven models**
locus of innovation	cultural transmission (acquisition)	cultural transmission (comprehension)	use (production)
cause of innovation	imperfect inference (poverty of stimulus)	imperfect inference (contextual mismatch)	extension in context (pragmatic plasticity)
realisation of innovation in	the hearer only	the hearer only	speaker and hearer

Table 2.1: Comparison of the three types of models of the cultural evolution of language.

treat the non-deviating, literal case as the limiting case rather than the norm, and state that "the hearer should take an utterance as fully literal only when nothing less than full literality will confirm the presumption of relevance. In general, looseness of expression is to be expected" (Sperber and Wilson 1995:234). I will come back to the connection between loose talk and relevance in chapter 3. The point to remember for the moment is that the key to understanding production-driven models is their view of *normal* language use as the extension of extant bits of code in concrete communicative contexts.

Table 2.1 summarises the differences between acquisition-driven, comprehension-driven and production-driven models with regard to where they locate innovation, what cause they assume for it and whether innovations are realised in both interlocutors involved in a communicative act or just in one of them. Acquisition-driven models and comprehension-driven models both refer to some sort of imperfect cultural transmission as the source of innovation. Consequently, innovations only happen in the hearer but not in the speaker. Production-driven models, in contrast, identify use as the locus of innovation. Consequently, innovations are realised in both interlocutors.

2.3.3 Which model to go for?

This is the point to return to the aims of this thesis. In chapter 1, I explained that it is the goal of this thesis to develop a mechanistic model of the cultural evolution of language that can account for (i) the *emergence puzzle*, that is, the fact that language emerged from no language, and (ii) the *design puzzle*, the fact that language exhibits the appearance of design for communication. As opposed to a nativist explanation, a culture-based approach to language evolution will try to identify the *minimal* set of *general*, non-language specific cognitive mechanisms that are *necessary* for the emergence and complexification of language. Since the discussed models are not mutually exclusive, it can easily be the case that two or more of the described processes have been involved in the evolution of language. Yet, some of the processes might characterise the cultural evolution of language—and the set of cognitive processes it minimally requires—more distinctively, while others might describe mere side-effects that have not substantially contributed to the evolution of language and without which language could still have evolved. We must thus judge which of the proposed models of the cultural evolution of language proves to be most promising in the light of the just mentioned aims. Here, the principles of *Ockham's razor* and the *uniformity of process*, as introduced in section 1.3.2, serve as evalution criteria. In the following paragraphs, I will point to some problems acquisition- and comprehension-driven models face in the light of the above requirements, and take the position that the production-driven model provides a promising alternative, whose potential has not yet been exploited enough in studies of the cultural evolution of language.

2.3.3.1 Problems with acquisition-driven models

As we have seen above, the core idea of acquisition-driven models of the cultural evolution of language has been adopted from the nativist generative framework of linguistics. Some of the problems that arise with such models can be traced back to this transfer of nativist conceptions to non-nativist explanations. Others are of a more general conceptual and empirical nature. I will discuss the following: (i) acquisition-driven models cannot account for the design puzzle, (ii) they do not provide an account for the emergence puzzle, (iii) they violate the principle of the uniformity of process, and (iv) some of these models make tacit assumptions that are questionable on the basis of empirical evidence. These are the assumptions that (i) there is a poverty of stimulus in language acquisition, and that (ii) grammars are non-redundant and language learners strive for grammar simplicity. Finally, I will argue that acquisition-driven models can be seen as a specific case of comprehension-driven models.

The design puzzle. One shortcoming of acquisition-driven models of the cultural evolution of language, which I have already mentioned in the introduction to this thesis, is that they fail to account for the design puzzle, that is, the appearance of complex adaptive design for communication reported for language. Researchers in the ILM-N framework in particular claim that their computer simulations corroborate the assumption that language adapts during the process of cultural evolution, but that this adaptation is to the learning bottleneck and not to the communicative function of language: language adapts to be learnable (see e.g. Brighton et al. 2005). Even though such models consider some sort of "function," and thus could be labelled as "functional" under certain circumstances (namely if we were to understand "function" as the pressure to which language adapts), they do not attribute much significance to the *communicative* function of language. Language, in these simulations, is often best described as an abstract system whose sole "function" is to be acquired through processes of grammar induction. But because acquisition-driven models do not consider communication as an important process in the cultural evolution of language—apart from its providing the public representations from which grammars are induced—they only show that cultural evolution can lead to *some* complex adaptive design (namely design for acquisition in their case—potential benefits for communication are seen as fortunate side-effects) but not how it leads to complex adaptive design *for communication*.

The emergence puzzle. Another shortcoming of acquisition-driven models is that they do not provide a story as to how linguistic communication emerged from a non-linguistic state. Language can only be *acquired* on the basis of language. Grammar induction alone cannot bridge the gap from an alingual state to one where linguistic communication is in place. The typical starting point for acquisition-driven accounts of language evolution is thus some sort of code, that is, a symbolic communication system consisting of *random* mappings of signals and meanings. But this means that the weight of explanation is moved to the internal—albeit coincidental—structure of these assumed initial states, as I have argued in Hoefler (2006b). At best, acquisition-driven models can be models of a *part* of the cultural evolution of language, namely of a process which kicked in once some primary symbolic system—some call such a system a "protolanguage" (but see Smith 2006c:321f. for a critical discussion of the usefulness of such a concept)—was already in place. Acquisition-driven models are thus the wrong tool to account for the emergence puzzle, namely to explain how language emerged from no language at all.

The uniformity of process. Further issues have been raised on the basis of empirical evidence from studies of both language acquisition and language change. Croft (2000:44–53) reviews these arguments against what he calls the "child-based theory of language change" in detail. He comes to the conclusion that the acquisition-driven view of the cultural evolution of language cannot be right because it makes four false predictions. The first prediction is that the "mistakes" that children make while they are acquiring language are of the same type as language changes which have actually occurred. Croft (2000:46f.) lists a number of empirical studies that show that this is not the case, that the "types of changes attested in historical language change are not identical to those found in language acquisition (Dressler 1974; Drachman 1978; Vihman 1980; Hooper 1980, Slobin 1997:313–14, Bybee and Slobin 1982:36–37, Aitchison 1991:168–73)." The second prediction an acquisition-driven model would make is that innovations or "errors" in children's grammars survive into adulthood. Croft (2000:47) holds against this prediction that it has been shown that "children's errors which presumably manifest a grammar (or lexicon) different from that of their parents tend to disappear in later phases of language acquisition." A third prediction is that language change would be abrupt since a previous generation would have an old version of the grammar whereas a new generation would have a new one. But Croft (2000:49–51) mentions that the propagation of innovations in a population is a gradual, social process, rather than a matter of mere generation turn-over. Finally, Croft (2000:51–53) points out that, as a fourth false prediction, acquisition-driven innovation would entail that a speaker would always ever have either the old or the new grammar at any one time. Change would be discrete. However, the phenomenon of "layering," the presence of an older and a younger form in a speaker's I-language as described, for example, by Hopper and Traugott (2003:124–26), renders this last prediction untenable too. Similar accounts of the arguments against child language acquisition as the locus of innovation are presented in Hopper and Traugott (2003:43–45) and in Traugott and Dasher (2005:41). Now, the principle of the uniformity of process would demand that if the acquisition-driven model is not valid as a model of language *change*, there is little ground to assume that it is a valid model of language *evolution*.

Tacit assumptions. Apart from the described conceptual and methodological problems which have been raised with regard to acquisition-driven models, there is also a set of tacit assumptions some of these models make which have to be met with certain reservations since they do not necessarily appear to comply with linguistic evidence. Most of these assumptions can be traced back to the generative origins of the model in the first place.

One of these assumptions is that there is a poverty of the stimulus in human language acquisition. This assumption provides one of the main arguments of linguistic nativism but, in the form of the so-called learning bottleneck, also constitutes a cornerstone of several acquisition-driven computer simulations. I refer the reader to section 1.1.1, where I have already reviewed the objections that have been raised against the assumption that there is a poverty of the stimulus.

Another tacit assumption often found in acquisition-driven models is that grammars are non-redundant, and that language learners strive for grammar simplicity. There is psycholinguistic evidence which suggests that this assumption cannot be maintained. Collocations (Barlow 2000) and idiomatic expressions indicate that linguistic knowledge is indeed stored redundantly: a concrete instantiation of a phrase can be part of the grammar along with the general rule that licences it. As any second language learner knows, it is often not enough to know the general rules of a language in order to produce acceptable sentences: there is also a large amount of usage-related conventionality involved. Goldberg (1995:ch. 8) discusses this by means of English resultive constructions: while *shoot someone dead* or *beat someone unconscious* are grammatical, **shoot someone wounded* or **beat someone dead* are not, even though there are no syntactic or semantic constraints that would prevent them from being grammatical just as well (see also Taylor 2002:554–58 for an overview).

Acquisition-driven models as comprehension-driven models. Acquisition-driven models and comprehension-driven models agree in two fundamental points: they both locate innovation in cultural transmission, and they both view the imperfect inference of the cultural artefact on the basis of observed use as the cause of innovation. Speaking in the terms of general cultural evolution, they both emphasise the variability of cultural transmission rather than its fidelity. We may thus ask if acquisition-driven models can be seen as a special case of comprehension-driven models. I will answer this question with "yes." The crucial point is to realise that the mechanisms by which innovation is introduced in the two models are essentially the same. In comprehension-driven models, innovation occurs because what the hearer infers is different from what the speaker had in mind due to a mismatch between the speaker's and the hearer's knowledge. In the acquisition-driven model, innovation occurs because the set of sample sentences from which the hearer infers her grammar, and the order in which he is exposed to them, is different from the set of sentences from which the speaker had inferred his grammar, and the order in which they had been presented to her. But such linguistic experience is just another sort of knowledge the speaker and hearer have. If it were the case that both speaker and hearer observe the same set of sample sentences in the same order, they would both infer the same grammar and no innovation would be introduced. The acquisition-driven model of the cultural evolution of language can thus be interpreted as a very specific case of a comprehension-driven model.

The conclusion of our considerations must be that acquisition-driven models of the cultural evolution of language do not seem to have the potential to meet the requirements of our project. We have seen that this is for the following reasons: (i) they cannot lead to complex adaptive design for communication, that is, they fail to account for the design puzzle, (ii) they do not provide an account of the emergence of language from no language, that is, they fail to explain the emergence puzzle, (iii) they violate the principle of the uniformity of process since empirical evidence discredits them as models of language change, (iv) they are built on a number of tacit assumptions which prove not to be psychologically plausible in the face of linguistic evidence, and finally (v) on a more abstract level, there is no need to consider them a distinct type of model because they are subsumed by the comprehension-driven model of the cultural evolution of language.

2.3.3.2 Comprehension or production: which is prior?

Burling (2000) proclaims the priority of comprehension over production: production-driven innovation would be a side-effect of comprehension-driven innovation. I will argue for the opposite. Burling (2000:27) says that "[c]omprehension runs consistently ahead of production" since we "are always able to understand more than we can say." The problem with this argument is that it ignores the fact that inferences are not only made on the part of the comprehender but also on the part of the communicator: speakers *invite* hearers to make specific inferences in addition to the recovery of the meaning *expressed* in the signal. Burling's claim that comprehenders can always understand more than communicators express must thus be complemented with the statement that communicators can always mean more than they express. And while it is true (a truism?) that for a communicator to complement what is expressed with inferences, she needs to assume that the comprehender is capable of understanding them, it is equally true that for a comprehender to complement what was expressed with inferences, he needs to assume that the communicator is capable of inviting them. Burling (2000) is right in stating that the capacity for comprehension must be in place when production occurs, but from an evolutionary perspective, the capacity for production must also be in place when comprehension occurs. This line of argumentation can thus not lead to answering the question of whether comprehension-driven models or production-driven models describe the more basic process—which is what we want to ask.

In many aspects, comprehension-driven models and production-driven models are more similar than a cursory look at the comparison in Table 2.1 would suggest. Both models are based on the observation that communicating does not only mean to encode and decode a signal but also to draw inferences from context. A communicated meaning M (the utterance or speaker meaning) is always a unification of the meanings recovered through the processes of decoding the produced linguistic form $\delta(F)$ and of inferring from context $\iota(C)$.[14] This is expressed in equation (7). Learning is then in both models a simple matter of what Kuteva (2001:6) refers to as "context-absorption": the addition of a new form-meaning pair $\langle F, M \rangle$ to one's linguistic knowledge G. This is expressed in equation (8), where t denotes the time before the learning event and $t + 1$ the time after it.[15]

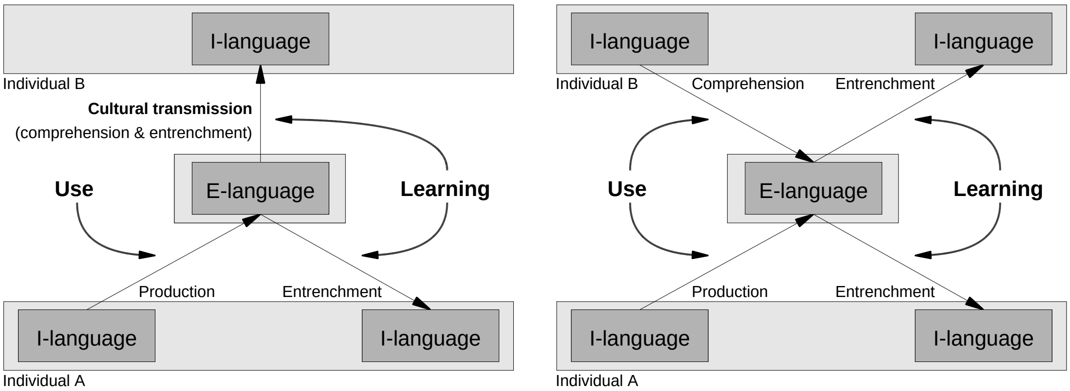
(7) Use: $M = \delta(F) \cup \iota(C)$

(8) Learning: $G_{t+1} = G_t \cup \{\langle F, M \rangle\}$

Comprehension mirrors production, and vice versa. Equation (7) applies to both production and comprehension. In production, the speaker "invites" the hearer to decode the signal and infer from context. And in comprehension, the hearer does what he thinks the speaker has invited him to do: he decodes the signal and infers from context. And in both the speaker and the hearer, the resulting usage of the produced linguistic form gets entrenched. The reason why this is not immediately evident from the descriptions given in Table 2.1 above lies in the way the three processes of *production*, *comprehension* and *entrenchment* are mapped onto the two processes of *use* and *learning*. This is visualised in Fig. 2.11. So far, we have grouped comprehension and entrenchment in the hearer together as one

[14] I will show later in chapters 3 that inference from context actually subsumes decoding.

[15] Note that the grammar is here described as an inventory or set of constructions (form-meaning pairs), and that learning is simply the incorporation of another construction, that is, the unification of the grammar with another set of constructions.



(a) Comprehension as part of learning (b) Comprehension as part of use

Figure 2.11: Comprehension mirrors production. (a) Comprehension and entrenchment in the hearer are conflated in the process of cultural learning. (b) Comprehension and entrenchment in the hearer are differentiated: comprehension mirrors production, both are processes of use. Learning in the form of entrenchment occurs in both the speaker (after production) and the hearer (after comprehension).

process of cultural transmission, that is, one of our two types of learning. This is represented in Fig. 2.11(a). However, just as we have split innovation (individual learning) into production and entrenchment, we can split cultural transmission (social learning) into comprehension and entrenchment. We thus get the more precise picture given in Fig. 2.11(b), which reflects the fact that comprehension mirrors production. Now we can map the two processes of the iterated learning model (see Fig. 2.6 above) more accurately: use can be either production or comprehension, and learning happens in the form of entrenchment in both the producer and the comprehender. This analysis is echoed in Croft's statement that "[e]ntrenchment is reinforced through use (comprehension as well as production)" (Croft 2000:73).

But then, where does the difference between comprehension-driven and production-driven models lie? While both are built on equations (7) and (8) above, comprehension-driven models are further constrained: they only consider cases where what the hearer infers is different from what the speaker invited him to infer due to a mismatch between the hearer's and the speaker's assumption about what constitutes context *C*. This additional constraint of comprehension-driven models is expressed in equation (9).

(9) $C_{Speaker} \neq C_{Hearer}$ (additional condition in comprehension-driven models)

	comprehension-driven models	production-driven models
use (production and comprehension)	coding and inference	coding and inference
learning	memorisation/entrenchment (“context-absorption”)	memorisation/entrenchment (“context-absorption”)
context	non-shared (common ground assumed)	shared or non-shared (common ground assumed)

Table 2.2: Comprehension- and production-driven models. The two aspects of use, production and comprehension, mirror each other: the speaker invites decoding and inference from context, the hearer decodes and infers from context according to what he thinks the speaker has invited him to do. The usage is memorised as a new form-meaning mapping, where the meaning is a unification of the decoded and the inferred meaning. The only difference between the two models is that comprehension-driven models additionally assume that there is a contextual mismatch so that the hearer's innovation might differ from the speaker's, whereas production-driven models are concerned with speaker-innovation and do not consider (but also do not deny) the potential second locus of innovation in the hearer. In both cases, however, speaker as well as hearer assume that there is shared context (common ground).

In this sense, comprehension-driven models are a specific case of production-driven models where a mismatch of context opens up a second locus of innovation. The process of innovation, however, is the same as for production-driven models: coding enhanced with invited (or assumed invited) inference from context and the subsequent process of context-absorption through entrenchment. For production-driven innovation to occur, it is enough to assume conditions (7) and (8) above. Invited inference and context-absorption are sufficient for production-driven models. Whether what the hearer infers is the same as what the speaker had in mind is not relevant as long as communication does not evidently fail. Comprehension-driven models also build on invited-inference and context-absorption—and thus on a situation which enables production-driven innovation. But they assume an additional condition, namely the one where speaker and hearer infer different things, and thus innovate differently. This comparison of comprehension- and production-driven models is summarised in Table 2.2.

To verify the assumption that production-driven models describe more basic processes than comprehension-driven models, we can try to prevent production-driven innovation from happening, and see if comprehension-driven innovation is still possible, and vice versa. To prevent comprehension-driven innovation we simply have to make sure that speaker and hearer have the same context. Production-driven innovation is not affected by this measure. On the other hand, if we try to prevent production-driven innovation, we have three possibilities. First, we can rule out invited inference. But if speakers did not invite inferences, hearers would have no reason to complement decoding with inference from context: even misinterpretations are based on the assumption that what is inferred is what the speaker meant. A second possibility to prevent production-driven innovation is to disallow context-absorption. But if we rule out this sort of learning, then the inferences made by the hearer, whether they are the same as the ones invited by the speaker or not, would not be able to enter the grammar. No innovation would happen at all. Thirdly, we could assume that the context is constant. This would also stop production-driven innovation from happening. But it would also mean that comprehension-driven reinterpretations would never become visible. Thus, comprehension-driven innovation can only occur under conditions that also allow for production-driven innovation, whereas production-driven innovations can occur whether comprehension-driven innovations are possible or not.

The bottom line is this. Both comprehension-driven and production-driven innovations are attested in language use and change. But the production-driven model describes the default case, whereas the comprehension-driven model describes an additional special case that can only occur under specific circumstances. In this case, it seems to me that before we resort to the exceptional case for explanations we should fully exploit the explanatory power of the default case. In the remainder of this thesis, I will thus develop and work with a production-driven model of the cultural evolution of language.

2.4 Conclusion

In this chapter, we have made a first step towards the development of a model of the cultural evolution of language which is capable of accommodating the puzzles of emergence and design. The starting point was a discussion of cultural evolution in general. In the course of this discussion, we have seen that human culture is special because it is cumulative. Humans engage in particularly faithful forms of social learning which allow them to accumulate modifications of artefacts (knowledge, skills, or social practice) beyond single individuals' lifetimes. However, the discussion has also shown that the driving force of cultural evolution, innovation, has been located at two different positions: in use or in transmission. The first view describes the process by which an individual "invents" or "improves" an artefact to make it fulfil a certain function in a given environment. The second view emphasises that, because social learning implies that an individual infers an artefact on the basis of an observed usage event (rather than acquiring it directly), alterations can be introduced during the process of cultural transmission. The first position thus conceives use as the driving force of cultural evolution: transmission merely ensures that innovations can accumulate beyond an individual's lifetime. For the second position, transmission itself is the innovative motor of cultural evolution.

These two perspectives of *general* cultural evolution are reflected in the production-driven and comprehension-driven models which have been proposed as explanations for the evolution and change of *language*. I have argued that, for conceptual and empirical reasons, these two types of models must be preferred to a third type, acquisition-driven models. A comparison of the two has shown that comprehension-driven models can be seen as an exceptional case of production-driven models: one in which the speaker and the hearer "misunderstand" each other. What both models assume (the comprehension-driven model tacitly, the production-driven model explicitly) is that speakers innovate in production by inviting hearers to complement the process of decoding with inference from context. Language use is intrinsically innovative because it exhibits pragmatic plasticity: the meaning linguistic forms come to convey in specific contexts go beyond the conventional meanings with which they have previously been associated.

When linguistic cultural evolution is described in terms of iterated learning, in fact, it includes two processes: learning (the mapping of E-language onto I-language) and use (the mapping of I-language onto E-language). Traditionally, studies in evolutionary linguistics have emphasised the innovative role of learning (hence, I guess, the term "iterated learning model"). But the process of comprehension, which, as I have suggested, mirrors production, has to be distinguished from the subsequent updating of the hearer's knowledge. In both production- and comprehension-driven models, innovation thus happens in use (production or comprehension) and not in learning (memorisation or entrenchment)—although such an innovation would, of course, be lost without its subsequent addition to the innovator's knowledge through learning. Maybe they ought to be called iterated usage models rather than iterated learning models?

What speaks against the employment of comprehension-driven models to study the origins of language is simply this: it seems to make sense to exploit the full explanatory potential of the default before one turns to the exceptional case in search of explanations. I claim that this has as yet not been done. The remainder of this thesis will therefore be dedicated to the development of a production-driven model of the cultural evolution of language to account for the evolutionary puzzles set out in the introduction. In the next chapter, I will work out the details of how pragmatic plasticity leads to linguistic innovation, what general cognitive mechanisms this involves, and how these mechanisms can be employed to address the emergence puzzle.

CHAPTER 3

The emergence puzzle

One of the puzzles that evolutionary linguistics has to explain is that of emergence: how can language have emerged from no language?[1] It is the question about the transition from a non-/pre-linguistic state to a state where linguistic communication is in place. This much seems to be agreed on by nativists and non-nativists: that any attempt to bridge this evolutionary gap must be able to account for at least the emergence of symbolism and the emergence of grammatical structure. Bickerton (2003:80), for instance, states that "language as we know it today involves the coming together of three things: modality, symbolism and structure." Even though he adds a third feature to the list, namely modality or the mode of transition (speech or signing), he then only mentions it very briefly, while he dwells much longer on the emergence of symbolism and structure. The central role he ascribes to these two is also reflected in the title of this essay: "Symbol and Structure: A Comprehensive Framework for Language Evolution." In the same volume (Christiansen and Kirby 2003), Tomasello, who, in contrast to Bickerton, is an outspoken opponent of the innateness assumption (see e.g. Tomasello 1995, 2004), makes the following statement:

> Human communication is most clearly distinguished from the communication of other primate species by its use of (1) symbols and (2) grammar. This means that progress on questions of language origins and evolution depends crucially on a proper understanding of these two phenomena. (Tomasello 2003b:94)

[1] Individual sections of this chapter appear in Hoefler and Smith (2008), Hoefler and Smith (in press). All material presented here is my own work.

However, the common assumption goes further than this. The emergence of symbolism and the emergence of grammar are generally viewed as two *distinct* explananda, phenomena with *different* origins.[2] Tomasello (2003b:109) summarises this common assumption when he argues that "different aspects of language—for example, symbols and grammar—may have involved different processes at different evolutionary times" and titles the corresponding essay "On the different origins of symbols and grammar." While both Tomasello and Bickerton agree that "[t]he most crucial thing to grasp about the emergence of symbolic representation is that it must have been primarily a cultural rather than a biological event" (Bickerton 2003:82) and that "there was no specific adaptation 'for' symbolic [...] communication" (Tomasello 2003b:109), they disagree over the emergence of grammar: Tomasello views it as the result of different cultural processes, syntacticisation and grammaticalisation, but Bickerton assumes an additional biological step. Like this, we can theoretically find any combination of biological and cultural explanations for the emergence of symbolism and the emergence of grammar: Bickerton's explanation is cultural for symbolism but biological for grammar, Tomasello invokes two separate cultural processes, Pinker (2003) seems to favour biological explanations for both.

The distinction of the emergence of symbolism and the emergence of grammar has also given rise to one of the field's most prominent debates: the dispute between advocates of a synthetic protolanguage (e.g. Bickerton 2000) and those who think that protolanguage was holistic (e.g. Wray 2000). The assumption underlying this debate—and at the same time the one point on which both sides agree—is that, in a first evolutionary step, symbolic communication emerged and was used as some sort of "protolanguage," and that only later, and in a different process, protolanguage developed into language through the emergence of grammar. What the two sides cannot agree on is the nature of protolanguage symbols and the process that led to the emergence of grammatical structure.

In this chapter, I question this common assumption that the emergence of symbolism and the emergence of grammatical structure are two distinct explananda, with different origins. I agree with Tomasello that both can be explained through processes of cultural evolution on the basis of general cognitive mechanisms. But while he identifies the cognitive processes that underlie the emergence of symbolism, he remains more or less agnostic as to the ones that lead to the emergence

[2]This is with the possible exception of Pinker (e.g. 2003), who uniformly ascribes the emergence of all features of language to natural selection for the transmission of propositional information.

of grammar. He merely points out that they lead to grammaticalisation and syntacticisation but that they have not been investigated enough. The following statement, which I have already cited in section 2.3, summarises the research that Tomasello thinks needs to be done:

> Exactly how grammaticalization and syntacticization happen in the concrete interactions of individual human beings and groups of human beings, and how these processes might relate to the other processes of sociogenesis by means of which human social interaction ratchets up the complexity of cultural artefacts, requires more psychologically based linguistic research into processes of linguistic communication and language change. (Tomasello 2003b:103)

I suggest in this chapter that if one follows Tomasello's call for more cognition-based research into processes of linguistic communication and language change, one comes to the conclusion that the *same* cognitive mechanisms that lead to the emergence of symbolism also lead to the emergence of grammatical structure, and that there is little ground to treat the two as distinct evolutionary explananda. I propose that these cognitive mechanisms, which form the *common* cognitive origin of symbolism and grammar, are the ones underlying human linguistic as well as non-linguistic communication to the present day. The remainder of this chapter is divided into four main parts. I begin with a discussion of the cognitive basis of communication (section 3.1) and then apply the identified psychological mechanisms to account for the emergence of symbolism (section 3.2) as well as to explain the emergence of grammatical structure (section 3.3). Finally, I discuss the consequences of the presented approach for the protolanguage debate (section 3.4) before I present my conclusions (section 3.5).

3.1 The cognitive basis of communication

Why should one study communication in order to explain the emergence of symbolism and grammar? In the last chapter, I have shown that cultural artefacts are shaped by the "ratchet effect" of cumulative cultural evolution. Cumulative cultural evolution is made possible because faithful transmission allows for the preservation of innovations; but the motor of cultural evolution is the process of innovation itself. I have argued that the type of innovation with the greatest explanatory potential is the one that comes about through use. Use becomes innovative when extant artefacts are applied in novel environments. Cultural artefacts are thus shaped in cumulative cultural evolution through iterated usage in ever new environments.

The way an artefact is used (its usage) is defined by (i) what it is used *for* (its function), and (ii) the conditions of the environment in which it is used. The usage of an artefact is thus the way in which it is applied to interact with the environment to achieve a goal, namely that artefact's function. And the function of an artefact can in turn be described by the change in the environment which the artefact is intended to bring about when used. In order to understand an artefact and how it emerged and evolved through cumulative cultural evolution, it is therefore essential to study its function and usage.

The artefact in question, the emergence of whose defining properties is to be studied, is language. The function of language is communication: language is used for communication. One therefore needs to ask what change in the environment communication brings about and how it can be achieved in individual usage situations. An understanding of the cultural evolution of language through iterated usage and its cumulative effects is thus inseparably linked to an understanding of what communication is and how language is used to achieve it. I will first discuss the general requirements that a model of communication has to fulfil, and then introduce the two currently predominant models of communication: the code model and the ostensive-inferential model of communication. I will conclude by arguing that these two do not represent two different modes of communication, but that the former is subsumed by the latter—a point which will prove to be an important key to explaining the emergence of symbols and grammar.

3.1.1 Communication: its goal and challenges

The aim of communication is (i) to *inform* a conspecific of some state of affairs, and (ii) to *elicit* a response to that newly gained information in that conspecific, to elicit a "perlocutionary effect" (Austin 1962). The function of language is thus ultimately a social one: to manipulate the behaviour of a conspecific. Communication is a way of achieving this by making some information available to the addressee which has not been manifest to him before, and which is assumed to trigger the desired response in him. The change in the environment that communication is intended to bring about is therefore first and foremost a cognitive one: a change in the cognitive state of the addressee. Communication can thus be said to involve a transfer of information from one individual's cognitive environment to another individual's cognitive environment.

But there is an obvious obstacle: humans are not telepathic. They cannot transmit information from one mind to another directly. Smith (e.g. 2003a) emphasises the relevance of this simple fact for our understanding of language evolution. He criticises models of language evolution that assume direct transfer. If humans were able to transfer information from one mind to another directly, they would not need to resort to signals, and their existence—for example in the form of language—would not make sense. We must conclude that a communication system like language is first and foremost a tool to tackle this problem: how can information be transmitted, and a conspecific's mental state be altered, if individuals do not have direct access to each other's mind?

The obvious—and sole conceivable—solution to this problem is for the communicator to modify the *physical* environment shared with the addressee in a way that (i) prompts and (ii) enables the addressee to recover the communicated information. Sperber and Wilson (1995) respectively speak of the communicator having to make manifest to the addressee (i) her communicative intention and (ii) her informative intention. Modifications in the physical environment can, for instance, be acoustic, as in the case of speech, or they can be visual, like in gestural communication, drawing or writing. Accordingly, Sperber and Wilson (1995:1) incorporate the modification of the physical environment and its role as an intermediary into their definition of communication:

> Communication is a process involving two information-processing devices. One device modifies the physical environment of the other. As a result, the second device constructs representations similar to representations already stored in the first device. Oral communication, for instance, is a modification by the speaker of the hearer's acoustic environment, as a result of which the hearer entertains thoughts similar to the speaker's own.

However, the question that still needs to be answered is *how*, by means of modifications of the physical environment, communication is actually achieved. I will therefore briefly introduce the two most commonly invoked accounts, which are generally referred to as the *code model of communication* and the *inferential (or ostensive-inferential) model of communication*. For each of the two models, I will (i) outline its core concept and then (ii) discuss how communication is achieved in that model.

3.1.2 Communication by coding

Especially in technical domains, the code model has proved to be an extremely successful interpretation of what is going on in communication. While it can be traced back to the thinking of Aristotle, the model has probably found its most seminal formulation in Shannon and Weaver (1949), which, at its time, laid the foundations of a whole new field, information theory[3], and continues to exert considerable influence on the development of linguistics. Despite some explicit field-internal criticism (e.g. Origgi and Sperber 2000; Smith 2005a; Hoefler 2006a), it is still widely applied in evolutionary linguistics.

The core element of the code model of communication is the *code*. A code is a predefined algorithm or system which is part of the interlocutors' knowledge and maps individual types of modifications of the physical environment (signals) onto meanings, and vice versa. It is a set of instructions that describe how meanings can be translated into signals and how signals can be translated into meanings. For communication to be established successfully, both the communicator and the addressee must have the same code at their disposal.

Communication is achieved by coding. The communicator uses her code to *encode* the information to be transferred into a signal, which she then produces.[4] By producing the signal she modifies the physical environment of the addressee. The addressee, upon perceiving this modification of his physical environment, uses his code in turn to *decode* the signal into the transferred information. "Communication is achieved by encoding a message, which cannot travel, into a signal, which can, and by decoding this signal at the receiving end" (Sperber and Wilson 1995:4). In the framework of the code model of communication, the type of physical environment chosen to transport the signal is usually referred to as the "channel." Communication by coding is summarised in Fig. 3.1. The communicator[5] intends to convey meaning A. She uses her code, which maps meaning A

[3]It needs to be pointed out, however, that with the advent of lossy data compression algorithms (cf. section 5.3.1), information theory has actually transcended the limitations of the code model of communication.

[4]It has become custom to refer to the communicator with feminine and to the addressee with masculine forms.

[5]In the context of the code model, the communicator is often referred to as the "sender," and the addressee as the "receiver."

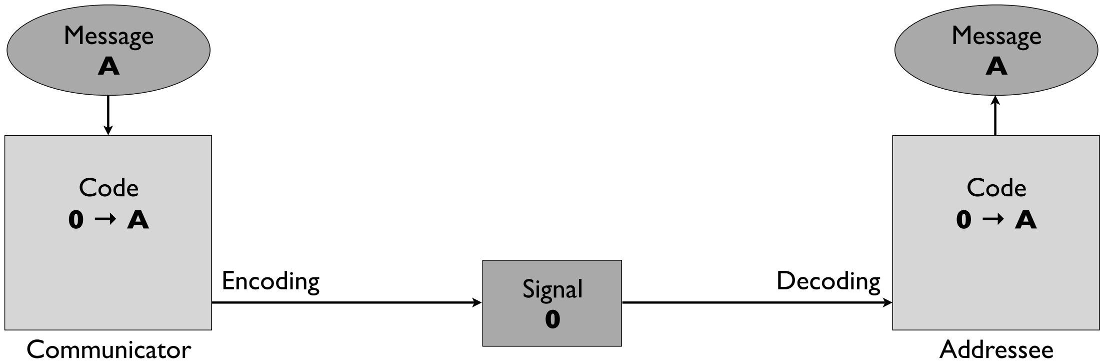


Figure 3.1: Communication by coding.

onto form *0*, to do so, and produces form *0* as a signal. The addressee perceives the signal and decodes it by using his own code to recover the communicated meaning *A*. Communication is thus achieved with the help of the code each interlocutor has, which is used to link meanings to specific modifications of the physical environments, the signal.

In coded communication, the signal fully specifies the communicated meaning, and a code ideally consists of one-to-one mappings of forms onto meanings. While *synonymy*, the case where one and the same meaning can be expressed by more than one form, is not so much a problem for successful communication—the communicator can simply choose one of them at random—the opposite case, *ambiguity*, represents a serious obstacle. If one and the same form can stand for more than one meaning, how can the addressee know which of these meanings the communicator had encoded when she produced the form? Without any additional mechanisms, that is to say, without transcending the code model, the addressee cannot solve the disambiguation problem.

But there are also two more general problems with the code model of communication. One of its shortcomings with respect to the study of language evolution is that it is symbolic. Symbolic association, namely the (relatively arbitrary) mapping of forms and meanings in a code, is its core element. It can therefore not offer any help if one attempts to explain how symbolism has come about in the first place. Adopting the code model of communication in studies of the origins of language may even lead to the misconception that, in order to get to language, humans needed to have a "symbolic capacity" as a cognitive prerequisite—a *specific*, innate module of the mind that associates forms with meaning. In constrast, I will argue in this chapter that symbolism is an emergent property, a by-product of iterated communicative interactions that is based on general cognitive capacities and cumulative cultural evolution. Another problem with the code model is, as has been pointed out frequently, that it does not accurately capture what is going on in linguistic communication: language users do not just encode and decode signals. Successful linguistic communication also heavily depends on inference from context. This fact is taken into account in the other major model of communication: the inferential model of communication.

3.1.3 Ostensive-inferential communication

The first articulation of the inferential model of communication is usually ascribed to Grice (1957, 1975). Just as Shannon and Weaver's (1949) discussion of the mathematical underpinnings of the code model of communication marked the birth of information theory, Grice's formulation of the inferential model stood at the beginning of a new field: pragmatics. And even though in recent years something of a division over certain linguistic specifics seems to have appeared between so-called neo-Gricean approaches on the one side (e.g. Levinson 2000; Horn 2004) and the relevance theorists on the other side (Sperber and Wilson 1995; Carston 2002), the inferential model of communication, in one form or another, is applied by all. Sperber and Wilson's (1995) investigation of the model—which they call the "ostensive-inferential model"—is maybe of most interest to evolutionary linguists since it attempts to abstract away from linguistic description to the underlying cognitive mechanisms.

I prefer Sperber and Wilson's name for the model because it mentions *both* processes which are involved in communication according to it: ostension and inference. In the ostensive-inferential model, "communication is achieved by producing and interpreting evidence" (Sperber and Wilson 1995:2). The former is ostension, the latter inference; the former is the communicator's part, the latter the addressee's. However, the concept at the core of the ostensive-inferential model, which is a prerequisite for establishing communication by producing and interpreting evidence, is that of context. The notion of context thus needs to be introduced and discussed before one is able to describe how communication is achieved in the ostensive-inferential model.

3.1.3.1 *What is context?*

The context of a communicative situation is the knowledge (or cognitive environment) which the interlocutors recognise as common ground in that situation. The first point to emphasise about context is therefore that it is *cognitive*. Often, the physical environment of a communicative situation is mentioned as an example of what context is. This is, however, misleading. Not the physical environment itself is part of the context of that communicative situation, but the communicator's and addressee's perception of it, or rather the knowledge that is derived from that perception. Context is always a part of the cognitive environment of the interlocutors at the time of the usage event.

However, context only comprises those parts of the interlocutors' cognitive environment which they recognise as *common ground* in the given situation. This means that context is knowledge which the interlocutors recognise as being shared in that situation. But the notion of common ground does not just imply (i) that they recognise the said knowledge as shared but also (ii) that they are aware that their counterpart recognises it as shared too—and (iii) that they realise that their counterpart also knows that they are aware of this. Sperber and Wilson (1995:88) refer to this as a "mutual cognitive environment." The concept of common ground has been addressed, for example, by Lewis (1969); Schiffer (1972); Stalnaker (1978); McCarthy (1990), but probably the most extensive discussion is provided by Clark (1996:92ff.).

Clark distinguishes between two categories of common-ground knowledge. *Communal common ground* is knowledge that an individual recognises as common ground because of the community to which its counterpart belongs. The two individuals assume that they share that knowledge because they are both members of, for instance, the same family, group, profession, nationality, or simply because they both belong to the same species and thus perceive the world in similar ways and share some basic experiences. *Personal common ground*, on the other side, is knowledge which two individuals recognise as being shared because it originates from a joint experience (present or past). This can, for instance, be their memory of events in the past at which they were both present, memories of their past interactions, or simply the awareness that they are currently in the same physical setting.

In any communicative situation, contextual knowledge draws from communal common ground as well as from personal common ground. The knowledge that we currently experience the same *physical environment* because we stand next to each other is part of our personal common ground. It has nothing to do with what communities we belong to. But the assumption that your perception of this physical environment is similar to mine originates from communal common ground: namely from the fact that we belong to the same species and therefore have similar perceptual and cognitive functions. The same is the case with the knowledge of the *social dynamics* of the situation we are attending to. That we both know what the goal of our present interaction is partly originates from our personal common ground: it depends on our joint experience of the conditions or events that led to it. However, to establish shared expectations with regard to the goal of our interaction, we must not only be aware of the specifics of the current situation but also of the general behavioural patterns of our species and the social conventions of our group. This latter knowledge stems from our communal common ground. These are just two examples of how contextual information about the immediate physical or social environment can build on a combination of personal and communal common ground.

Common ground knowledge can amount to a shared understanding of the goal of a given interaction, or at least to an agreement of what plausible goals for it are. Two individuals can only establish and recognise common ground on the goal of an interaction if they also share a relatively detailed understanding of each other's intentions (cf. Tomasello 2003a). Grice (1975) describes this cooperative goal-directedness of interactions for the case of talk:[6]

> Our talk exchanges do not normally consist of a succession of disconnected remarks, and would not be rational if they did. They are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a *common purpose* or set of purposes, or at least a *mutually accepted direction*. The purpose or direction may be fixed from the start (e.g., by an initial proposal of a question for discussion), or it may evolve during the exchange; it

[6]One of the differences between Gricean pragmatics and Relevance Theory is that the former puts more emphasis on cooperation whereas the latter focuses more on the self-interest of the individual. Sperber and Wilson (1995:161) distinguish Grice's approach from theirs in the following way: "Grice assumes that communication involves a greater degree of cooperation than we do. For us, the only purpose that a genuine communicator and a willing audience have in common is to achieve uptake: that is, to have the communicator's informative intention recognised by the audience." However, at the level of abstraction that I assume here, the intricacies of this distinction do not come to bear.

> may be fairly definite, or it may be so indefinite as to leave very considerable latitude to the participants (as casual conversation). But at each stage, some possible conversational moves would be excluded as conversationally unsuitable. (Grice 1975:45, emphases added)

From such shared understanding of the interactional goal also falls common-ground knowledge of what is *relevant* in the context of a specific interaction. We can define relevant information as information that (i) is not accessible to the addressee yet, and that (ii) would contribute to achieving a goal that appears plausible in the given interaction if it was transferred to him. Information is relevant if communicating it would elicit an effect in the addressee that matches an interactional goal which makes sense in the given situation. The notion of relevance is at the center of the cognitive approach to communication developed by Sperber and Wilson (1995). However, they work with a somewhat idiosyncratic definition of relevance (see e.g. Bach 2006:7 for some criticism). For the present purpose, it will suffice to state that two individuals who engage in communicative interaction discern relevant from irrelevant information on the basis of knowledge which they recognise as common ground.[7]

3.1.3.2 How is communication achieved?

In the ostensive-inferential model, communication is achieved by producing and interpreting evidence. By modifying the immediate physical environment, the communicator (i) adds another fact to the common ground, which, in the given context, (ii) triggers the inference of the communicated information in the addressee. Communication is thus established through an *ostensive* act performed by the communicator (the modification of the physical environment by means of which she enhances the common ground) and an *inferential* act performed by the addressee but predicted and invited by the communicator (the inference of some new information on the basis of the now altered context). If, for instance, I want you to know that I have a one pound coin in my pocket, I can simply take the coin out, show it to you and make sure that you see how I put it back into my pocket. This is the ostensive part of communication. You will then infer from

[7]Note that both the knowledge of common ground in general and the knowledge of what constitutes relevant information in particular are themselves inferred from more basic knowledge about the world, one's conspecics and the current situation: "[f]or the speaker [the communicative context] includes information that could reasonably be inferred about the addressee, which gives rise to expectations about how the addressee will react to certain kinds of linguistic stimulus and certain bits of information" (Wedgwood 2007:651).

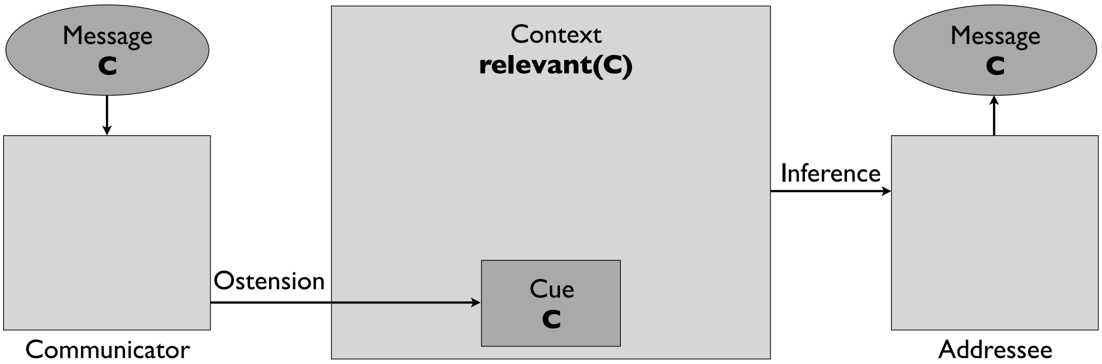


Figure 3.2: Ostensive-inferential communication by producing direct evidence.

what you have just seen that I have a one pound coin in my pocket. This is the inferential part of communication.

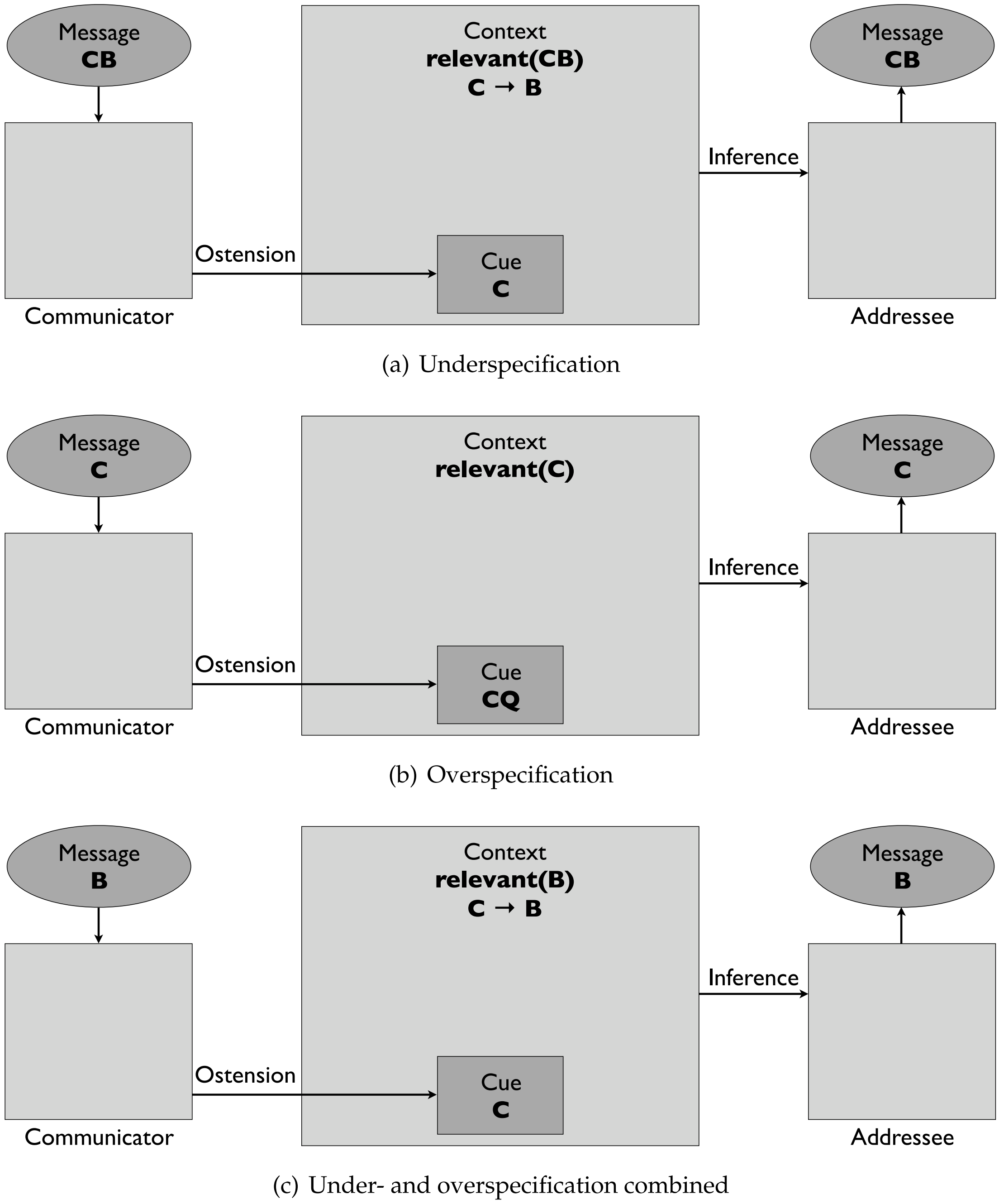
In the simplest case, the communicator provides *direct evidence* for the information she wants to transfer. This is a maximisation of the ostensive part and a minimisation of the inferential part: I show you the coin in my hand (the ostensive part), and therefore you know that there is a coin in my hand (the inferential part). Ostensive-inferential communication by means of direct evidence is represented schematically in Fig. 3.2. A communicator wants to transfer information *C* (e.g. that there is a coin). He produces direct evidence for *C*, such that *C* becomes part of the addressee's cognitive environment too. Because the addressee perceives the communicator's ostensive act, he acquires information *C* and, since *C* consitutes relevant information in the given context, infers that this is the information that the communicator wanted to convey. In this simplest case of ostensive-inferential communication, the produced evidence itself is the transferred information.[8]

[8]With regard to the introduced schematic representation, remember that the item that signifies knowledge of what constitutes relevant information in the given situation is itself a shortcut for a range of inferences drawn by the interlocutors. On a similar note, it needs to be mentioned that one thing that is always to be inferred from the context is the communicator's communicative intention: the fact that the behaviour she exhibits is communicative and therefore comes with the presumption of relevance. The communicative character of an observed behaviour, its "signalhood," can be inferred if that behaviour does not otherwise make sense in light of the expectations one has about the other individual's motives and intentions in the given situation. Even though it is not represented explicitly, the inference of the communicative character of the produced behaviour, and the recognition of this fact as common ground, is presupposed in all examples.

The common ground of two interlocutors often contains background knowledge related to the evidence produced. But as long as this knowledge does not allow the inference of some relevant additional information, it will be ignored. We might, for instance, both know that one pound coins are issued by the Bank of England, that all one pound coins issued in the same year have the same design, or that having several of them can make one's purse annoyingly heavy. But as long as none of this knowledge allows you to infer information that is relevant to our current interaction, you will ignore it.

On the other hand, provision of full evidence is often not feasible. In such situations, communicators can rely on shared contextual knowledge related to some producible evidence. The production of the evidence then functions as a *cue* to trigger the inference of additional information from the given context. If, for instance, I want to let you know that I have the one pound coin I need for my bus, I can show this to you by means of direct evidence when we both enter the bus at the same time and you observe how I pay my ticket with that coin. However, if we are nowhere near the bus yet (but on our way there), the production of direct evidence is not an option. In this case, I can still produce the coin as a cue and rely on our common ground: (i) that we have the same assumption of what information is relevant in the current situation, and (ii) that we both know that one pound coins are used to pay for the bus. Thus, if I produce a one pound coin while we are on our way to the bus station, and show it to you, you will not only infer the direct information contained in the cue, namely that I have a one pound coin, but, since this information alone would not be relevant in the given situation, you also infer additionally that this coin is for the bus.

Hence, whereas in communication by means of producing direct evidence, the cue fully specifies the conveyed information, in other situations, cues *underspecify* what is being communicated and trigger the inference of additional information relevant in the given context. Underspecification in ostensive-inferential communication is represented in Fig. 3.3(a). A communicator wants to transfer information *CB* (e.g. that he has a coin, and that this coin is for the bus). Her ostensive act adds *C* to the addressee's cognitive environment. And they are both aware that the knowledge that *C* is associated with *B* (that one pound coins are used to pay for the bus) is part of their recognised common ground. Because the new information only becomes relevant if *B* is added, the addressee infers that what the communicator means to convey is information *CB*. The produced cue *C* thus only underspecifies the intended meaning *CB* but can still be used to convey the



(a) Underspecification

(b) Overspecification

(c) Under- and overspecification combined

Figure 3.3: Under- and overspecified ostensive-inferential communication.

full meaning because its addition to the given context triggers the inference of the remainder of the intended meaning. The addressee will follow the associations he recognises as common ground until he arrives at a relevant interpretation. Carston (2004:822) describes the procedure employed by the addressee accordingly: "[c]heck interpretive hypotheses in order of their accessibility—that is, follow the path of least effort until an interpretation that satisfies the expectation of relevance is found; then stop."

Conversely, not all information contained in a cue is necessarily relevant: a cue *overspecifies* the actually communicated information if it has aspects which are ignored because they are irrelevant in the given context. The coin that I show you to let you know that I have a one pound coin for instance also has the property of carrying an effigy of the Queen. However, as long as this information is not relevant in the current situation, you will not consider it as part of what I mean to tell you. Overspecification in ostensive-inferential communication is represented schematically in Fig. 3.3(b). The communicator produces a cue CQ (e.g. a one-pound coin carrying an effigy of the Queen) to convey information C alone. The produced cue is thus more specified than the intended meaning: it overspecifies the actually communicated meaning. The communicator assumes that the addressee will infer from the context that he can ignore the aspect Q of the cue (e.g. the information that the coin carries an effigy of the Queen) because Q is irrelevant in the given situation.

In most cases of ostensive-inferential communication, underspecification and overspecification occur together. If, for instance, I show you my one pound coin to let you know that I have my coin for the bus, then the produced cue underspecifies *and* overspecifies the communicated meaning: in addition to the information that I have a coin, you infer from the context that this coin is for the bus (underspecification). But at the same time, you ignore certain properties of the produced cue because they are irrelevant: that the coin carries an effigy of the Queen, maybe, or that I hold it in my right rather than my left hand (overspecification). To some degree, every case of ostensive-inferential communication contains elements of under- and overspecification: the addressee always at least infers that the communicator wants to let him know something (i.e. the communicator's communicative intention), and always ignores certain specific features of the produced cue.

In the limiting case, a cue can be *maximally* under- and overspecified: when none of the information contained in the cue is part of the communicated meaning (maximal overspecification), and thus all relevant information is inferred from the context once the cue has been produced (maximal underspecification). If you ask me if I am going to town by bus or by bike, and I produce a coin in response, then I use a cue in such a way. You will infer that I intend to go by bus on the basis of our common ground knowledge that one pound coins are used to pay for the bus, and because they are not associated in any way with bikes. At the same time, the fact that I have a coin as such is not relevant in the present situation: all I want to communicate (and all you want to know) is that I am going by bus—whether or not I have the money to pay for it does not matter in this interaction. This situation of maximal under- and overspecification is represented schematically in Fig. 3.3(c).

Cues in ostensive-inferential communication thus exhibit *pragmatic plasticity* because they can under- and/or overspecify the conveyed meaning. Like this, the information that is actually transferred in a particular communicative situation can deviate from the information that the communicator immediately adds to the common ground by producing a cue. I conclude that pragmatic plasticity—comprising both under- as well as overspecification—is a fundamental property of human ostensive-inferential communication and not specific to language.

3.1.4 Coding as an ostensive-inferential process

It has been argued, for example by Sperber and Wilson (1995), that the code model of communication and the ostensive-inferential model of communication simply describe two *modes* of communication applied by humans, and that while, for instance, human ad-hoc gestural communication is ostensive-inferential, linguistic communication makes use of both communication by coding and ostensive-inferential communication. Language use, in this view, has a coding aspect and an inferential aspect.

In such a model of linguistic communication, the output of a coding process serves as the input to an inferential process. A speaker encodes a meaning into a linguistic signal, and the hearer decodes that signal to recover the encoded meaning. The meaning encoded in the linguistic signal then additionally serves as a cue to trigger further inferences from the context. This view of linguistic communication as a combination of the code model and the ostensive-inferential model is represented in Fig. 3.4(a). A speaker intends to express meaning *AB*. She uses her linguistic code to translate *A* into a linguistic form *0*, which she then

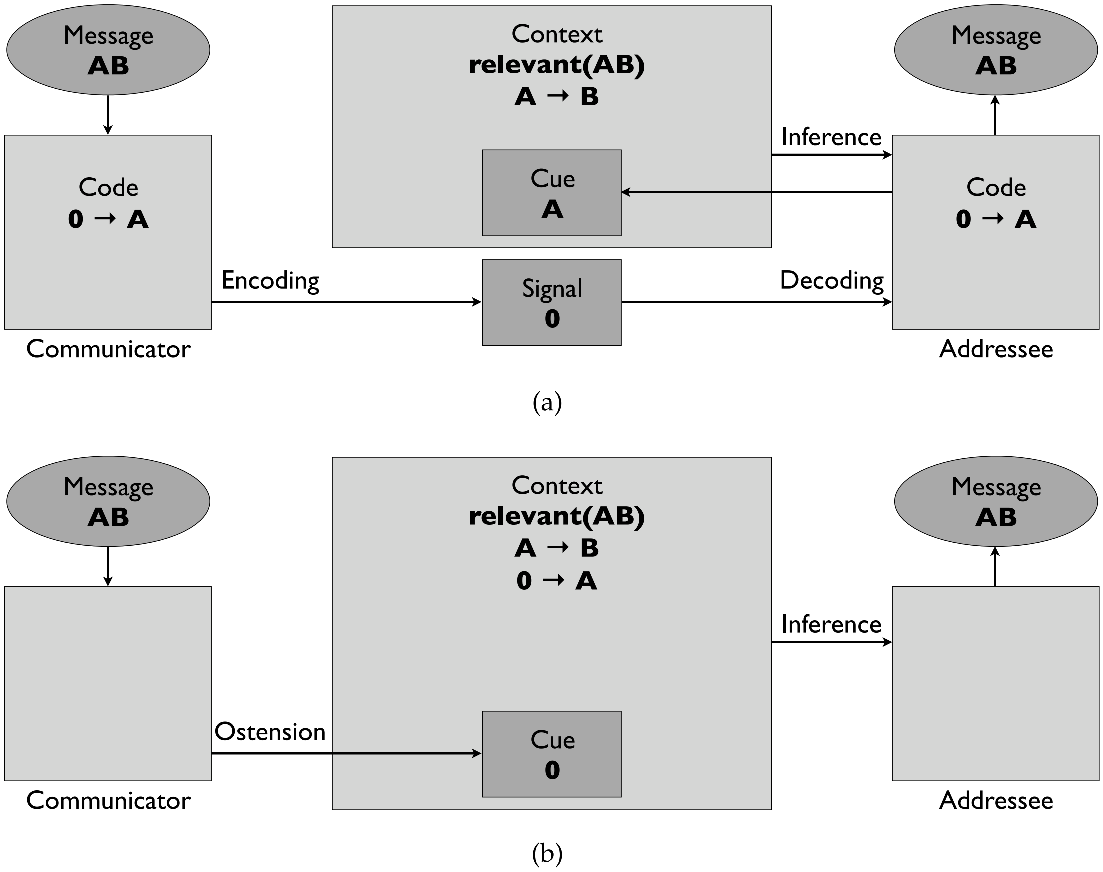


Figure 3.4: Linguistic communication. (a) The output of the coding process as the input to the inferential process. (b) Coding as an ostensive-inferential process.

produces. The hearer decodes the linguistic signal he has perceived and thus recovers *A* as the meaning encoded in the signal. Like this, the coding process has added new information, *A*, to the context. Once *A* is added to this particular context, it triggers the additional inference of *B* and thus allows the hearer to recover the full meaning intended by the speaker, *AB*.

While this description of linguistic communication may be useful for certain purposes and at a certain level of abstraction, it creates the illusion that there are two cognitively different processes involved: coding and inference. Sperber and Wilson (1995), for instance, conclude that linguistic communication has to be seen as having two aspects, coding and inference, because the processes employed in coding are fundamentally different from those used in ostensive-inferential communication:

> Inferential and decoding processes are quite different. An *inferential process* starts from a set of premises and results in a set of conclusions which follow logically, or are at least warranted by, the premises. A *decoding process* starts from a signal and results in the recovery of a message which is associated to the signal by an underlying code. In general, conclusions are not associated to their premises by a code, and signals do not warrant the messages they convey. (Sperber and Wilson 1995:12f., emphases in the original)

I argue that it becomes evident that this dichotomy is not warranted once one dissolves the borderline between code and context. A linguistic code, at least according to the view held in Cognitive Linguistics, consists of a set of individual linguistic conventions (constructions, each associating a form with a meaning) which can be paraphrased as the knowledge that "conventionally, form *f* is produced to communicate meaning *m*" and that therefore "when *f* is produced, *m* is meant." But then, there is no reason to assume that the knowledge of an individual that a linguistic form *f* is used to convey a meaning *m* is *in principle* different from the knowledge of that same individual that a one pound coin is used to pay for the bus.

Like other knowledge too, linguistic conventions can be part of the personal or communal common ground that two individuals share. And while linguistic conventions are part of two individuals' communal common ground, it is not the case that the only reason why two individuals share linguistic conventions is that they belong to the same language community. Individuals share linguistic conventions because of their membership to all kinds of communities, for instance because they have the same profession, or are of similar age. While two Koreans, being part of the same language community, share to a large extent the particular set of grammatical constructions generally identified as the Korean language, they might not share all lexical constructions if one of the two is fifteen years old and the other is eighty, or if one is an ophthalmologist and the other is not. A Korean ophthalmologist and an English ophthalmologist, on the other hand, might share a fair number of words but might not share any of their grammatical constructions, unless the Korean also has knowledge of English—or vice-versa. Clark (1996:116), citing Hopper et al. (1981), shows how very idiosyncratic linguistic conventions can even be maintained as personal common ground by giving examples from personal idioms that couples use among themselves.

If linguistic conventions are recognised as common ground knowledge and thus as part of the context of a communicative situation like any other knowledge too, then coding appears as a normal ostensive-inferential process, namely one where linguistic forms are used as cues to trigger the inference of the meanings conventionally associated with them. Fig. 3.4(b) exemplifies this. The speaker intends to convey information *AB*. Both interlocutors recognise as common ground the linguistic convention that form *0* is used to express meaning *A*, and furthermore that *A* can entail *B*. By producing the signal *0*, the speaker adds a new fact to this common ground which then in turn triggers the inference of first *A* and then *B* in the given context. At the same time, the hearer realises that the fact that *0* was produced itself does not contain any relevant information. Communication by coding is thus a special case of ostensive-inferential communication with maximally under- and overspecified cues that trigger inferences against the background of linguistic conventions contained in the context.

The argument that Sperber and Wilson (1995:27) use against the interpretation of coding as an ostensive-inferential process is that "[a] variety of species, from bees to humans have codes which are to a greater or lesser extent genetically determined." This, of course, is the innateness assumption which this thesis exactly does *not* want to make, as explained in the introduction. Sperber and Wilson (1995:173f.) advocate a Chomskyan version of it. If we drop the innateness assumption, no reason remains why human coding should not be seen as a case of ostensive-inferential communication. Sperber and Wilson (1995:26) themselves concede that the use of non-innate codes—which they refer to as "artificial codes" to distinguish them from their innate cousins—is an ostensive-inferential process:

> Regard a code as a set of conventions (in the sense of Lewis 1969) shared by all participants in the communication process. Members of the audience use their knowledge of these conventions on the one hand, and their knowledge of the signal and of the context on the other, to infer the message. This is a reasonably good description of what often happens when artificial codes are devised and used.

My preliminary conclusions are therefore these. Linguistic communication is ostensive-inferential. Like any sort of ostensive-inferential communication, it is achieved by producing and interpreting evidence against the background of recognised common ground which, among other things, includes the shared idea of what is relevant in the given situation. And like in all forms of ostensive-inferential communication, its cues can under- and overspecify the actually conveyed information. This is consistent with the view of linguistic communication advocated in Cognitive Linguistics, as expressed in the already mentioned statement by Kemmer and Barlow (2000:xxi, emphasis in the original):

> The context-dependent nature of linguistic production and understanding entails, among other things, the inevitable underspecification of linguistic forms. Language does not hold or "convey" meaning per se, but simply provides *cues* for meaning construction in context.

What distinguishes linguistic communication from other types of ostensive-inferential communication is simply an additional kind of common-ground knowledge it makes use of: conventional associations between linguistic forms and meanings. The conclusion must therefore be that the ostensive-inferential model not only describes instances of human ad hoc non-linguistic communication but is also sufficient to account for linguistic communication. It is a model of all human intentional communication.

We are thus now in a position where we can further specify the evolutionary step from no language to language as a step from non-linguistic ostensive-inferential communication to linguistic ostensive-inferential communication. The question that needs to be answered is then where the form-meaning mappings used in linguistic ostensive-inferential communication come from. This is ultimately the question about the emergence of symbolism.

3.2 The emergence of symbolism

Symbolism is a phenomenon so definitive of human language that it lends itself to be viewed as the result of one evolutionary step: the biological emergence of a "symbolic capacity" from a pre-symbolic state (which is often not specified any further). I have chosen a different avenue. I will argue that symbolism can emerge gradually from ostensive-inferential communication without any further biological modification, and that the set of general cognitive capacities required for ostensive-inferential communication is sufficient to explain its emergence. I thus agree with, for instance, Tomasello (2003b) or Bickerton (2003), who also assume that symbolism is the product of cultural evolution.

The key to explaining the emergence of symbolism, in my opinion, is to break the phenomenon down into more elementary components, and then to explain the emergence of each of these components individually. To account for the emergence of symbolism then means to answer these two questions: (i) how do forms become associated with meanings, and (ii) how do form-meaning associations become arbitrary?

3.2.1 How do forms become associated with meanings?

The knowledge of a linguistic convention, which associates a form with a meaning, is not in principle different from any other sort of knowledge. To stick to our example, the knowledge that a form *f* is conventionally used to convey a meaning *m* is not fundamentally different from the knowledge that one pound coins are used to pay for the bus. The first question to answer, if one intends to explain the emergence of symbolism, is thus how such knowledge is acquired—how it is learned.

Learning is a cognitive change based on *experience*. Somebody learns that one pound coins are used for the bus either because another person tells them, or because they experience it themselves when they have to pay for the bus for the first time—or when they observe someone else doing it. Knowledge, in this latter case, comes about through usage. Similarly, an association of a form with a meaning can be described as the experience of what that form was used for in the past: that when form *f* was produced as a cue, it was used to communicate meaning *m*. The knowledge of a form-meaning association is thus ultimately also experience-based.

Forms become associated with meaning through the *memorisation* of their *usage*. When a communicator successfully uses a form *f* as a cue to convey a meaning *m*, both the communicator and the addressee will remember this usage event: they will remember that the production of *f* was used to trigger the inference of meaning *m*—similarly to the way in which our passengers will remember that they used a one pound coin to pay for the bus. At a very basic level, form-meaning associations can thus be interpreted as the memory of previous usage events.[9]

[9]The type of learning, which I here describe as memorisation of usage events is refered to as "exemplar-based learning" or "instance-based learning" in the machine learning literature (e.g. Mitchell 1997). I will explain how the usage-based model of language evolution introduced in this chapter maps onto these approaches to learning in chapter 4.

The memorisation of previous usage can have two effects: (i) entrenchment and (ii) the establishment of new common-ground knowledge between the two interlocutors and their consecutive ability to refer to it in future usage events. *Entrenchment* is the term used in the context of usage-based approaches to language (e.g Langacker 1987, 2000; Barlow 2000; Tomasello 2003a) to refer to the process of automatisation or routinisation that occurs when a form is repeatedly used in the same way. Entrenchment as such is not specific to communicative behaviour alone: "[e]ntrenchment simply refers to the fact that when an organism does something in the same way successfully enough times, that way of doing it becomes habitual...." (Tomasello 2003a:300). Langacker (1987:57–60) describes how composite forms and composite meanings which are sufficiently entrenched become psychological units, and how, in the same way, form-meaning associations gain unit-status.

A sufficiently entrenched association between a form and a meaning can thus be accessed without the complicated reasoning that was involved when the association was not established yet. As an example, one can think of charades-like games, where one person has to communicate a concept to a group of other people only by means of pantomimic gestures. When the presenter has to convey a particular concept for the first time, she may have to do a fair amount of reasoning to come up with a gestural description that would allow her audience to infer the right meaning. And likewise, the audience would have to invest some reasoning to figure out what the presenter's gestures are meant to denote. But when the same presenter, or another person from the same group of people, faces the task of conveying the same concept for the second time, she does not need to do all the reasoning again but can simply re-use the gestures which were successful the previous time. And the audience, upon her re-using the said gestures, will not have to do all the guesswork that is involved in figuring out the meaning of a new gesture, but will be able to understand more readily what the presenter means on the basis of their memory of the previous usage event for those gestures. Such re-usage under the (implicit or unconscious) assumption that what worked before will work again leads to automatisation.

Usage-based studies of language (e.g. Barlow and Kemmer 2000) often stress the relation between entrenchment and *frequency* of use: the more frequently a form f is used to express a meaning m, the more deeply the association between the two will be entrenched. And the deeper a form-meaning association is entrenched, the more readily it is accessible. Each usage of f to express m thus increases the depth of entrenchment of the respective form-meaning association in the knowledge of both the communicator and the addressee. To repeat Croft's comment, "[e]ntrenchment is reinforced through use (comprehension as well as production), and decays through lack of use (as any rusty second language learner can attest)" (Croft 2000:73).

Once the association between a form *f* and a meaning *m* is sufficiently entrenched, form *f* can be used to convey meaning *m independently* of any situation-specific context. The inference of *m* will now be triggered immediately by the production of *f* and will not have to be constructed on the basis of additional situation-specific common-ground knowledge anymore. Through automatisation, forms can thus get to express meanings independent of the original context: pragmatically inferred meaning becomes conventional ("semantic") meaning.

But the conventionalisation of form-meaning associations is also due to another consequence of the memorisation of the form being used to convey the meaning: every usage event creates new common ground between the involved individuals. When two individuals communicate, they establish shared experience. After two individuals have used a form *f* to communicate a meaning *m* for the first time, the association of *f* with *m* through that usage event becomes part of their personal common ground. But once the experience of a new co-occurrence of a form and a meaning has become part of common ground knowledge of two individuals, this experience can be referred to in future communicative situations and thus itself serve as a context from which novel meanings are inferred.

The more entrenched the association between a form and a meaning is, the more easily it can be used as part of the context of future usage situations: invoking the memory of a shared one-time usage event involves more explicit reasoning than employing an established association between a form and a meaning. Conventionalisation is thus the result of a combination of both entrenchment, on the one side, and the creation of new common ground which can be referred to in future situations, on the other side. Conventionalisation is a matter of degree: an association becomes more conventionalised the more frequently it is referred to, and also the more individuals are involved. However, what is minimally required for a form-meaning association to become a convention is only (i) that the form has been used to convey the meaning in at least one usage event, and (ii) that this usage event has become part of the personal common ground of at least two individuals.

The cognitive prerequisites for the conventionalisation of form-meaning associations are thus the same as the ones for ostensive-inferential communication. Like ostensive-inferential communication, conventionalisation is based on the recognition of common ground, in particular on the recognition of personal common ground knowledge of shared previous experience. The additional requirement for conventionalisation is rather a *social* one: two individuals have to communicate the same meaning to each other more than once—otherwise conventionalisation cannot be brought to bear.

The described processes, which ultimately lead to the establishment of a new convention, can be observed, for instance, in the experiments documented in Garrod et al. (2007). In these experiments, "pictionary" tasks are used to study the evolution of graphical signs. Pairs of participants are asked to repeatedly communicate a limited set of concepts to each other by means of graphical representations. Garrod et al.'s main finding, and its information-theoretic interpretation, is that conventionalisation happens as a "shift in the locus of information from the sign to the users' memory of the sign's usage" (Garrod et al. 2007:961). While in the beginning, the burden of reasoning is on the produced cue and its interaction with the *situation-specific* context, later on it moves to the cue and the memory of its previous usages. The first time the participants have to communicate a certain concept, they use iconic representations to do so. These same representations are re-used later when they have to convey the same concept again. In these later rounds of the experiment, the meaning of the drawing is then identified more quickly because not only the original context (the set of all occurring concepts plus general cultural knowledge) but also the history of earlier usages of the signs is available. This "shortcut" is also reflected in the evolution of the form of the re-used signs: the drawings used in later phases of the experiment do not have to resemble the concept anymore but simply need to identify a previous usage of that sign and can therefore be much more abstract than their original. That the participants refer to their shared memory of previous usages can be recognised in situations where the meaning of some drawing is identified before the drawing is completed—in a way in which the audience could not have identified the meaning without memory of previous drawings and the recognition that one of them is being repeated.

3.2.2 *How do form-meaning associations become arbitrary?*

Symbols are often distinguished from other types of form-meaning associations by the arbitrariness of the relationship between their forms and their meanings. Saussure (1916) views this arbitrariness between what he calls the signifier and the signified as the defining characteristic of (linguistic) signs. Peirce (1932) distinguishes three types of relationships between forms and meanings. The relationship between form and meaning is iconic if the form resembles the meaning perceptually or structurally. This is the case, for instance, for the pictograms used at train stations or airports: a schematic representation of a telephone receiver points to a public phone. The relationship between form and meaning is indexical if it is causal: smoke is an indexical sign for fire. In symbols, finally, the relationship between form and meaning is arbitrary.[10] Hockett (1960) lists the arbitrariness of linguistic signs as one of the design features of human language. Hurford (2003:48) also mentions the arbitrariness of linguistic form-meaning associations as a defining characteristic of human language and concludes that there must be a connection between the fact that linguistic form-meaning associations are arbitrary and the fact that the associations are learnt rather than genetically inherited.

Therefore, if one attempts to account for the origin of symbols, it is necessary that one asks how *arbitrary* form-meaning associations can emerge. So far, I have shown how non-arbitrary form-meaning associations originate in ostensive-inferential communication. But the relation between a cue and the information whose inference it triggers in a given context cannot be arbitrary.[11] The question about the emergence of arbitrary form-meaning associations can now be approached in two ways. Either one assumes that arbitrary and non-arbitrary form-meaning associations emerge independently of each other: ostensive-inferential communication leads to the emergence of non-arbitrary form-meaning associations, and another process, cultural or biological, has to be found to account for the emergence of arbitrary form-meaning assoications; or, and this is the avenue that I will take, one can ask how non-arbitrary form meaning associations become arbitrary.

[10] Peirce (1932) also points out the conventionality of symbols. But this does not clearly distinguish symbols from icons: which icon is used for a specific meaning can be a matter of convention.

[11] An arguable exception is situations where the addressee is capable of inferring what the communicator wants to say by virtue of the fact alone that the communicator has tried to catch the addressee's attention. In such cases, the cue employed to catch the addressee's attention need not be related to the communicated meaning.

There are two general ways in which the relation between a form and the meaning associated with it can become arbitrary: either the features of the form change, or the features of the meaning change. Even though in reality both occur, it is theoretically sufficient to have only one of these two types of changes to get from non-arbitrary to arbitrary form-meaning associations, and thus to be able to account for the emergence of symbolism. Examples of the former can be found, for instance, in the abovementioned experiments carried out by Garrod et al. (2007), where the form of originally iconic signs become more abstract in the course of iterated use, which finally leads to a situation where the original iconicity of the form is not recognisable anymore. Sound changes as described by historical linguistics, or the graphical change in the characters of writing systems can have the same effect. In this thesis, however, I will focus on the second option: semantic change. I will show that semantic change can be explained by the same set of cognitive mechanisms that I have applied so far to account for the emergence of form-meaning associations from ostensive-inferential communication. Of all the mentioned avenues to explain the emergence of arbitrary form-meaning mappings, this is therefore the most parsimonious one since it does not employ separate cognitive processes.[12]

Semantic change occurs when an extant form-meaning association exhibits pragmatic plasticity in use, that is, when it is used as a cue in an under- or overspecified way so that the actually communicated meaning deviates from the meaning conventionally associated with the produced form. When this novel usage of the said form enters common ground and is conventionalised, *polysemy* results: the interlocutors' linguistic knowledge now contains two mappings for the same form, one that associates it with the old meaning, and another one, which associates it with the new one. This situation is equivalent to what historical linguistics calls *layering*. If the new mapping is reinforced through frequent use whereas the entrenchment of the old one decays because it is hardly used anymore (and maybe eventually even drops out of use), the result is that the form involved has effectively changed the meaning associated with it.

The conventionalisation of underspecified and overspecified use have opposite effects. In the case of underspecification, entrenchment leads to a new form-meaning association with the same form but a more specific meaning, because the actually communicated meaning contains not only the meaning previously

[12]In fact, as I suggest in section 5.3.1, changes of form may be explained by means of the very same cognitive processes too.

conventionally associated with the form but also additonal meaning which was inferred from the context. In the case of overspecification, on the other hand, entrenchment leads to a new form-meaning association with the same form but a less specific meaning. Since some information contained in the meaning that is conventionally associated with the form is irrelevant in the given context, the actually communicated meaning contains less information than the form's original conventional meaning. The entrenchment of underspecified use thus brings about semantic *narrowing* and *specialisation*, whereas the entrenchment of overspecified use results in semantic *broadening* and *generalisation*.

The different effects of conventionalised under- and overspecified usages, and how they can result in the emergence of arbitrary form-meaning associations, can be illustrated with the following artificial example, which is summarised in Table 3.1. Imagine a situation where an individual intends to inform her conspecifics that a lion is around. This individual could do so by imitating a lion's roar. The addressees would infer from the context that the communicator's roaring only makes sense if it is interpreted as a communicative cue. They also recognise as common ground that lions roar and thus conclude that the communicator intends to convey information about a lion by means of the produced cue. The convention which can emerge from this under- and overspecified use of roaring is one which captures that roaring is used to convey information about a lion. The relationship between its form and its meaning is at the same time indexical and iconic: roaring is caused by lions, but roaring also perceptually resembles lions, or at least their acoustically perceivable side. In Table 3.1, this is represented as usage event 1.

Once it has emerged, the new convention might then be used in a context, where it is not only relevant that a lion is around but also that that lion poses a danger in this situation (for instance, because it is obviously hungry, or because it is a mother accompanied by cubs). After such a usage, the convention will become entrenched with an additional semantic aspect, that of danger, and sooner or later, roaring will only be imitated when a *dangerous* lion is around. This is represented as usage event 2. In some usage situations, the fact that there is a lion might not be considered terribly relevant by the addressees: what is relevant is that there is danger. After usage event 3, the convention's aspect of danger will thus be enforced while its lion aspect becomes less salient and eventually disappears (usage event 3). Once the group has moved away from lion territory, they may start to use roaring to warn of new dangers, for example snakes,

usage event	communicated meaning	situational context	produced cue	type of use	emerging convention
1	L	$R \rightarrow L$ $relevant(L)$	R	underspecified (overspecified)	$R \rightarrow L$
2	DL	$R \rightarrow L$ $L \rightarrow D$ $relevant(DL)$	R	underspecified	$R \rightarrow DL$
3	D	$R \rightarrow DL$ $relevant(D)$	R	overspecified	$R \rightarrow D$
4	DS	$R \rightarrow D$ $D \rightarrow S$ $relevant(DS)$	R	underspecified	$R \rightarrow DS$
5	S	$R \rightarrow DS$ $relevant(S)$	R	overspecified	$R \rightarrow S$
6	BS	$R \rightarrow S$ $S \rightarrow B$ $relevant(BS)$	R	underspecified	$R \rightarrow BS$
7	B	$R \rightarrow BS$ $relevant(B)$	R	overspecified	$R \rightarrow B$

Table 3.1: Arbitrariness emerges through iterated usage-induced semantic change. Each row represents a usage event in the example scenario which introduces a minimal semantic change because a cue is used (and entrenched) in an under- or overspecified way. Legend (refer to the description of the chain of events in the text): R = roaring, L = lion, D = danger, S = snake, B = belt.

and the convention might acquire a new ‘snake’ aspect. This is shown as usage event 4. However, if the group moves to a region with harmless snakes only, the ‘snake’ aspect of the convention might be strenghened while the ‘danger’ aspect is deemed irrelevant, as shown in usage event 5. Finally, snake skin might be manufactured into belts by this group of humans, so that the form used to denote snakes might also be used in an underspecified way to denote snake belts (usage event 6). And if the group moves on once more, to areas where snakes are rare, and the term for ‘snake’ is only ever used to denote belts anymore, the convention might again lose one of its aspects and come to denote any sort of belt, whether made of snake skin or not. This last step is represented as usage event 7 in the table. After many communicative situations in which extant conventions were used as cues in an underspecified or overspecified way, the relationship between the form and its meaning has eventually become arbitrary: there is no obvious connection between roaring and a belt. However, form-meaning associations are arbitrary only from a *synchronic* perspective; if considered from a *diachronic* perspective, they can theoretically be traced back to initial non-arbitrary[13] associations which have emerged from non-conventional ostensive-inferential communciation (Keller 1998).[14]

We can conclude that the emergence of symbolism is based on the same cognitive capacities that underlie ostensive-inferential communication. Form-meaning associations emerge when individuals memorise that a specific form was used to communicate a specific meaning. Form-meaning associations become conventional when this memory enters the common ground of at least two individuals. Once this has happened, the new associations can themselves be used as part of a context in a novel usage situation. Their forms can be used to communicate meanings that are more specific than the conventional meanings associated with them (underspecified use), or more general (overspecified use). The entrenchment and conventionalisation of these novel usages leads to semantic narrowing in the case of underspecification and semantic broadening in the case of overspecification. The relationship between forms and meanings becomes arbitrary

[13]By “non-arbitrary,” I mean here that, in a specific situation, the meaning of a produced cue can be inferred on the basis of the properties of the cue and general common-ground knowledge only, i.e. without the knowledge of a conventional form-meaning association.

[14]Note that this is not to say that all such chains of change need to start with an iconic form-meaning pair as their starting point. The presented example has only been chosen to illustrate the processes at work.

in the course of the cultural evolution of conventions due to such iterated usage-induced semantic change. The emergence of symbolism can therefore be accounted for by virtue of only (i) the cognitive mechanisms involved in ostensive-inferential communication, and (ii) the ability to remember usage events and to recognise and use them as common ground in later usage situations. For this last point, a minimum of two individuals is required that *repeatedly* indulges in communication. But do these general cognitive prerequisites also suffice to explain the emergence of grammatical structure?

3.3 The emergence of grammatical structure

Cognitive Linguistics has brought forward (or rather revived) a fundamental tenet which, I think, must at least cast some doubt on the common assumption that the emergence of symbolism and the emergence of grammar are two distinct evolutionary explananda: the so-called *symbolic thesis*. The symbolic thesis says that the linguistic competence of an individual can be described *exhaustively* as an inventory of form-meaning associations (e.g. Langacker 1987; Taylor 2002; Evans and Green 2006). I have already introduced this view of grammar in 2.2.3.2 above. The term "construction" is commonly employed to refer to linguistic form-meaning associations—or "conventional linguistic units," as Langacker (1987:57) calls them—and grammatical theories based on the symbolic thesis are thus generally referred to as construction grammars.

So why does the symbolic thesis pose a potential problem to the common assumption that symbolism and grammar are different explananda? It is because of one of its consequences, namely that one must assume that lexicon and syntax form a continuum. As opposed to generative grammar, where meaningful lexical items are organised by inherently meaningless syntactic rules, construction grammars hold that syntactic schemata are meaning-bearing constructions too (e.g. Goldberg 1995). The symbolic thesis and the lexicon-syntax continuum that falls out of it therefore lead to the following preliminary hypothesis: since, like lexical items, syntactic constructions are symbolic, they have emerged like all other symbols too. If this hypothesis is right, then it would seem that the cognitive capacities employed so far to explain the emergence of symbolism would be sufficient to account for the emergence of grammar too.

Syntactic constructions have two properties: they are abstract schematic patterns and they express very general, so-called "functional" meanings. Consider, as an example, the English passive construction: the pattern [X *be* VERB-*ed by* Y] signifies that X is affected by Y in a manner specified by the VERB. This construction is an abstract pattern (represented by the order in which its components are arranged) and has a very broad and general meaning (that one entity is affected by another entity). One thus has to answer two questions with respect to the emergence of syntactic constructions: (i) how do schematic constructions emerge, and (ii) how do extant constructions, whether schematic or specific, acquire the sort of broad, general meaning that has traditionally—and somewhat unfortunately since the terms are ambiguous—been referred to as "grammatical" or "functional"?

Both phenomena are usually described as processes of *grammaticalisation*—some (e.g. Tomasello 2003b) distinguish between syntacticisation and grammaticalisation, where the former describes the emergence of schematic constructions and the latter how meanings become more functional. However, it is important to note that grammaticalisation is just that: a linguistic description or classification rather than an explanation. It describes what changed into what, rather then why and how this change has occurred. The term "grammaticalisation" merely subsumes a particular set of phenomena of linguistic change where linguistic material loses its independence of use and becomes more functional. In order to *explain* phenomena of grammaticalisation, the underlying cognitive mechanisms need to be studied, as has been pointed out, for instance, by Heine (1997); Kuteva (2001); Tomasello (2003b).

Grammaticalisation is often accompanied by both semantic and phonological change; for the purposes of this chapter, however, only the semantic change involved is relevant. Phonological change is a mere result of grammaticalisation, but neither its prerequisite nor its cause. There is rather a causal chain that leads from semantic change to phonological change. The type of semantic change found in grammaticalisation leads to a change in the frequency and independence of use, and this change in the frequency and independence of use in turn brings about phonological change. This interpretation of grammaticalisation is supported by linguistic evidence, which shows that (i) in the process of grammaticalisation, phonological change occurs after semantic change, and that (ii) grammaticalisation can occur without phonological change. The English phrase *gonna*, for example, must have developed from *going to* only after *going to* had changed its meaning from expressing physical motion to expressing future and intention. This is evident because, like the contracted form, the non-contracted form *going to* can also be used with a future sense, but (in most English dialects) the contracted form *gonna* cannot be used to express physical motion (Trask 2000:141). This fact is illustrated in examples (1) and (2):

(1) a. I am going to be in trouble.
    b. I am gonna be in trouble.

(2) a. I am going to the beach.
    b. $*$ She is gonna the beach.

The initial hypothesis, that syntactic constructions emerge like all other symbols too, can now be refined. The cognitive mechanisms underlying the emergence of symbols, and also responsible for the semantic change symbols undergo subsequently, are those involved in ostensive-inferential communication. We can thus hypothesise that these same mechanisms account for (i) syntacticisation, that is, the emergence of schematic constructions, and for (ii) grammaticalisation, namely the process through which constructions become more functional. In the following sections, I will sketch the general properties of an approach to the emergence of grammar that is based on these hypotheses.

3.3.1 How do schematic constructions emerge?

The emergence of schematic constructions is not any different from the emergence of other form-meaning associations if one views form as what it is: information. Depending on the number of dimensions of the chosen physical channel, the information contained in a produced cue can be of various forms. It can, for instance, be the information that a specific sound or gesture was produced. But it can equally well be the information that two sounds or gestures—or two whole cues—have been produced in a specific order. The fact that one cue was produced after another cue can itself serve as a cue in a given context and trigger the inference of additional meaning. Such cues, and the information they communicate, can become entrenched and conventionalised, like any other form-meaning pair too.

Some basic word order can, for instance, emerge as follows. Imagine a situation where an individual intends to communicate to a conspecific some information Y about some state of affairs X. One discourse strategy the communicator can choose in this situation is to draw the addressee's attention to X first and therefore to let him know what the current exchange will be *about*. Once this is done, the cue for Y can be added. The addressee will thus not only infer X and Y but also that Y in some way concerns X. In this hypothetical case, the order in which X and Y are produced serves as a cue to communicate the theme and the rheme of the exchange.

This new form-meaning association may be used frequently in situations where the communicator says about an agent what action they perform. The theme-rheme aspect might be less relevant in such situations than the information that X is the agent and Y is the action performed by this agent. This new usage may become entrenched and conventionalised itself and, after more generalisation through usage-induced semantic change, lead to a syntactic construction which expresses a subject-predicate relation.

This view of the emergence of schematic constructions is compatible with other accounts of the emergence of syntactic patterns from discourse strategies. One of the most seminal foundational studies in this respect is presented by Givón (1979), who suggests a path of grammaticalisation not only from discourse to syntax but also from syntax to morphology, from morphology to morphophonemics and from there to zero. His famous quote that "[t]oday's morphology is yesterday's syntax" (Givón 1971:413) can easily be reformulated to announce that "today's syntax is yesterday's discourse." A similarly prominent case for the emergence of grammatical structure from discourse has been made by Hopper (1987). Tomasello (2003a) shows how children acquire schematic constructions from discourse, and concludes that something similar must have happened in language evolution. Accounts of the emergence of syntactic patterns from discourse strategies are frequent in the functionalist linguistics and grammaticalisation theory literature. These approaches view more complex syntactic phenomena such as long-distance dependencies or constraints on "movement" as *emergent properties*, that is, as patterns that arise during the cumulative cultural evolution of langauge out of the accretion of conventionalised schematic construtions and the way their forms and meanings interact. I agree with this view but also acknowledge that proof of concept, for instance in the form of a computational model that simulates the emergence of such more complex syntactic properties, has not been provided so far.

For the current purposes, the crucial thing to note is that the emergence of syntactic patterns from discourse strategies can be conceived as a special case of the emergence of a convention that associates a cue, which was produced in a specific context, with the meaning whose inference it triggered in that context. As is the case for other form-meaning associations, schematic constructions too can become arbitrary after their meaning has changed sufficiently because of iterated under- and overspecified use. The emergence of schematic constructions too is thus the result of ostensive-inferential communication and the memorisation of previous usages. Syntacticisation is based on the same cognitive capacities as the emergence of symbolism.

3.3.2 How do meanings become more functional?

Broadly speaking, two types of approaches to grammaticalisation can be identified, corresponding to speaker-driven models and hearer-driven models of language change respectively (see 2.3). The former type of explanation emphasises the role of metaphor and metonymy in grammaticalisation (e.g. Heine et al. 1991), the latter views reanalysis as its core aspect (e.g. Hopper and Traugott 2003). I present a unified perspective, which can be reached if one abstracts away from linguistic description to the underlying cognitive mechanisms. I will argue (i) that both what I will call the "metaphor-based scenario" and the "reanalysis-based scenario" can have occurred, but also (ii) that the two are not fundamentally different but simply represent variants of normal usage-induced semantic change that merely differ due to the specific conditions of the usage event in which they were introduced.

The example I use to illustrate the common psychological underpinnings of the two scenarios is that of the aforementioned grammaticalisation of the English construction *be going to*, which is one of (if not *the*) most cited and best documented examples in the grammaticalisation literature (see e.g. Heine et al. 1991; Kuteva 2001; Hopper and Traugott 2003), and also represents a particular instance of grammaticalisation which is very common, both historically and cross-linguistically (Heine and Kuteva 2002). The construction *be going to* has undergone a development from a state where it simply denoted spatial motion to one where it came to express intentionality and future even in situations where physical motion is not implied. These two meanings of *be going to* are illustrated in sentences (3) and (4):

(3) Where are you travelling to? I am going to London.
(*spatial motion*)

(4) What are your plans for tonight? I am going to stay at home.
(*intention/future* without *spatial motion*)

3.3.2.1 *Metaphor-based scenarios*

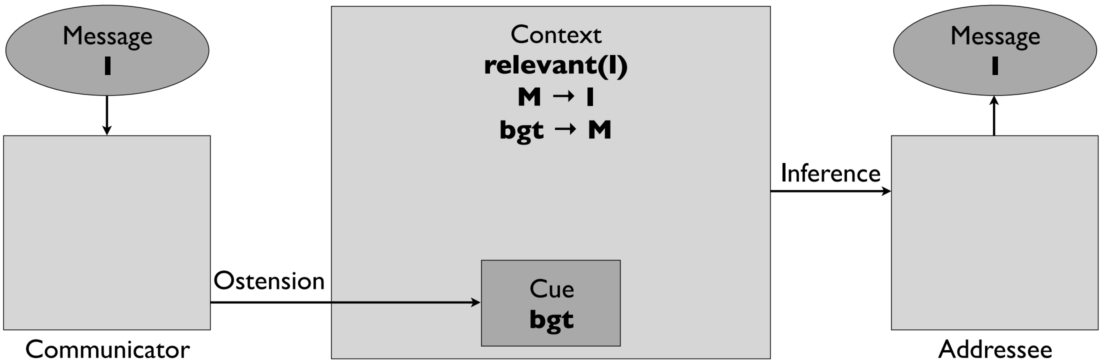
Metaphors[15] overspecify the meaning they communicate in a usage situation: they come to stand for another concept with which they conventionally share some but not all properties. In this sense, metaphors are at one end of a continuum representing the various ways in which conventional meanings can overspecify utterance meanings. In metaphors, the difference between the conventional meaning and the utterance meaning is strikingly big—or, to use Lakoff and Johnson's (1980) terminology, the distance between the source and the target domain is particularly large. The view that metaphors, along with other tropes, simply represent points on a continuum of "normal" overspecified language use is consistent with the approach to figurative language advocated by Sperber and Wilson (1995:231–37) and later refined in Carston (1997), and in Wilson and Carston (2007). The relevance theorists refer to overspecification as "loose" language use, and claim that literal use and metaphorical use form a continuum and only differ in their degree of looseness, with various other types of figurative language ranging somewhere in between. This perspective of metaphor is also consistent with the accounts of figurative language given e.g by Langacker (1987:68–71) or Croft (2000:99–114).

In what I call the metaphor-based scenario, the process of the grammaticalisation is thus initiated in a communicative situation where the speaker uses the linguistic convention which associates *be going to* with spatial motion in an overspecified, metaphorical way to express intention. This is only possible in situations where both interlocutors recognise as common ground (i) that spatial motion usually involves intention, and that (ii) in the given context, information about spatial motion itself is irrelevant. This scenario is illustrated in Fig. 3.5(a), and the speaker's and the hearer's reasoning is detailed in example (5).[16]

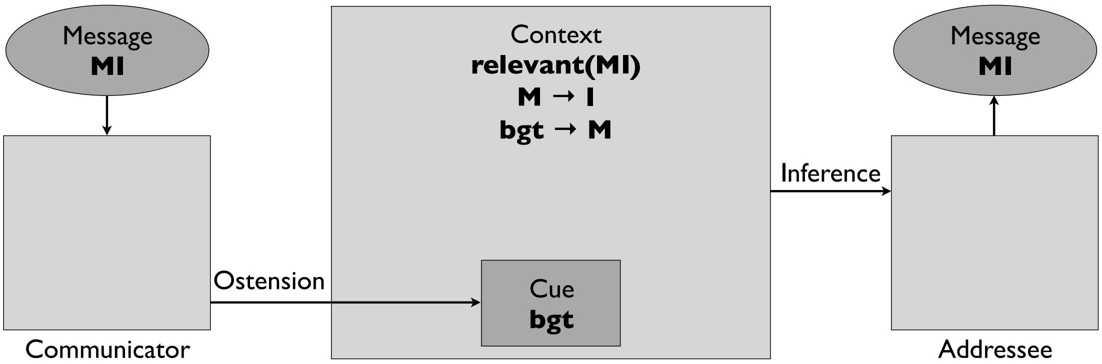
(5) Speaker's reasoning:
  a. I want to express 'intention'.
  b. I have a construction which expresses 'spatial motion', and the hearer shares this convention.
  c. 'Spatial motion' is associated with 'intention'.

[15]It needs to be emphasised that I am referring to novel ad hoc metaphorical language use here, and not to established and conventionalised metaphors.

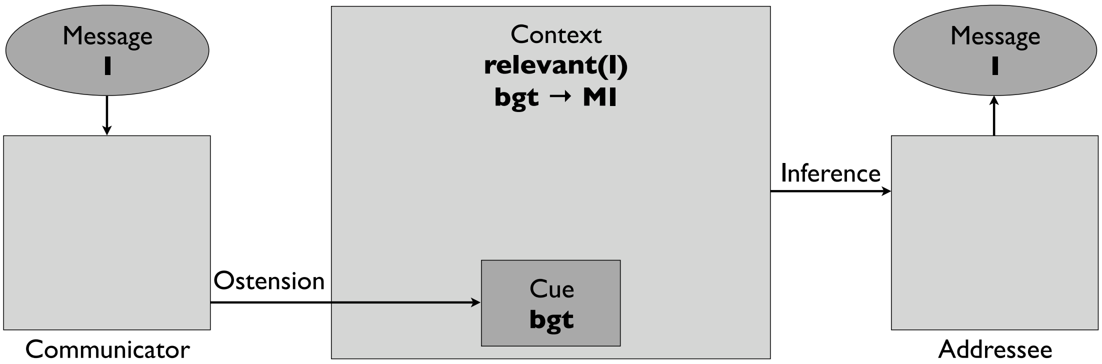
[16]Note that I have explicated the interlocutors' reasoning for explanatory reasons. I do not claim that these are very conscious processes; in fact, it must be assumed that the degree to which such reasoning is conscious depends on the degree of entrenchment of the involved associations: more explicit reasoning is involved in less familiar situations.



(a)



(b)



(c)

Figure 3.5: Metaphor-based scenarios of grammaticalisation of *be going to*. Legend: *I* = 'intention', *M* = 'spatial motion', *bgt* = *be going to*. (a) The metaphor-based scenario in one step. (b)–(c) The metaphor-based scenario in two steps.

d. ‘Spatial motion’ is not relevant in the given context.
e. Because we share common ground, the hearer will be aware of (b)–(d), and realise that I am aware of it too.
f. Because of (e), I can use the construction for ‘spatial motion’ metaphorically to convey ‘intention’.

Hearer’s reasoning:

h. The speaker has expressed ‘spatial motion’.
i. ‘Spatial motion’ is not relevant in the given context.
j. ‘Spatial motion’ often implies ‘intention’.
k. ‘Intention’ would be relevant in the given context.
l. I must assume that the speaker is cooperative.
m. I must also assume that the speaker is aware that I know (h)–(l), and that I know of his being aware of it.
n. From (h)–(m), I conclude that the speaker intends to convey ‘intention’.

The metaphor-based scenario can also be thought of as having an intermediate step. In the scenario as it is represented in Fig. 3.5(a), a new association that maps *be going to* onto the meaning ‘intention’ is created after only one usage event: after the described interaction, both speaker and hearer will remember that *be going to* was used to communicate intention. However, a similar scenario can be thought of that leads to the same result in two steps, as illustrated in Figs. 3.5(b) and 3.5(c). The starting point is also a state where *be going to* is only associated with spatial motion. In a first usage event, shown in Fig. 3.5(b), it is used in an underspecified way to express spatial motion but at the same time also intention. This is possible if the two interlocutors recognise as common ground that (i) spatial motion can imply intention, and that (ii) both are relevant in the present situation: the information that spatial motion happens and the information that intention is implied. After this usage event, *be going to* will newly be associated with a combined meaning of ‘spatial motion’ and ‘intention’. In a second usage event, shown in Fig. 3.5(c), this new form-meaning association is then in turn itself re-used as contextual information. Now, the speaker wants to express intention in a context where spatial motion is clearly not relevant. She can thus produce *be going to* in an overspecified way, correctly assuming that the hearer will ignore its aspect of ‘spatial motion’. Now, the two interlocutors will entrench a new form-meaning association that pairs *be going to* with the meaning ‘intention’ alone. The same result as in the first scenario has thus been reached, the only difference being that it was initiated by two rather than just one usage event.

What distinguishes the one-step scenario from the two-step scenario is not so much the cognitive mechanisms at work—both are based on underspecified and overspecified usage—but the degree of metaphoricity involved. The two-step scenario assumes two usage events where the difference between the meaning conventionally associated with the cue and the meaning actually communicated is only small, whereas the one-step scenario is based on a situation with a much more conspicuously metaphorical usage, where the conventional meaning and the communicated meaning do not overlap at all. Both scenarios are possible, and which of them really *has* occurred is not relevant for our purposes since, in terms of the cognitive processes they describe, they are essentially identical.

3.3.2.2 Reanalysis-based scenarios

The concept of reanalysis has been defined in a variety of ways in the different contexts of syntactic change, semantic change, and grammaticalisation (e.g. Timberlake 1977; Langacker 1977; Heine et al. 1991; Harris and Campbell 1995; Haspelmath 1998; Croft 2000; Campbell 2001; Detges and Waltereit 2002; McDaniels 2003; Hopper and Traugott 2003). Based on Langacker (1977:58), reanalysis has often been defined as a change in the underlying structure without any visible change in the surface structure. Such a general formulation, of course, is purely descriptive: it merely states the product of a change by comparing the state before and the state after the change. Depending on how the notion of "underlying structure" was interpreted, the definition has been applied to shifts in morpheme boundaries (e.g. Langacker 1977), changes in syntactic structures (e.g. Harris and Campbell 1995; but see Detges and Waltereit 2002 for an argument that reanalysis is not primarily syntactic but semantic), and generally to the reinterpreation of a pragmatically inferred meaning as a semantically encoded meaning—and, possibly, vice-versa. This last sense, of course, merely describes the result of conventionalisation, as can be seen in Traugott and Dasher (2005:35), who account for "the conventionalizing of pragmatic meanings and their reanalysis as semantic meanings." By this descriptive definition alone, any conventionalisation of under- or overspecified usages can be seen as an instance of reanalysis.

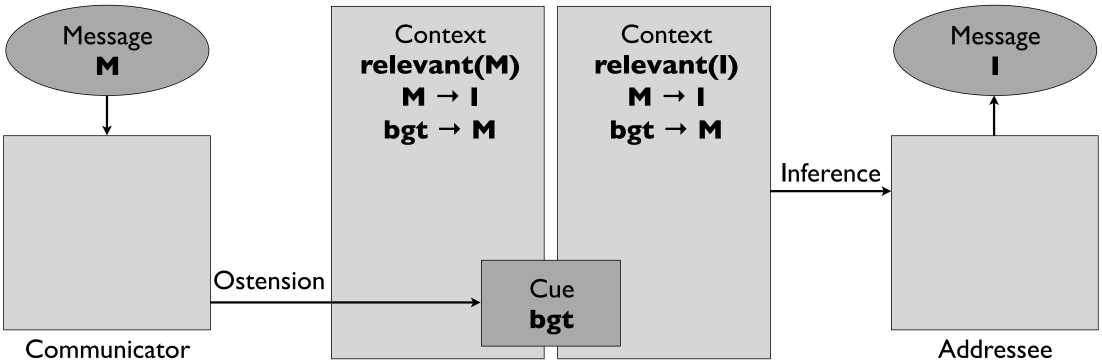
Often, however, reanalysis as a *process* is further constrained: it is described as happening under the condition of a mismatch between what the speaker assumes to be common ground in a given situation, and what the hearer assumes to be common ground in that situation (e.g. Croft 2000; Kuteva 2001; Detges and Waltereit 2002; Smith 2006a). Reanalysis-based scenarios of the grammaticalisation of *be going to* build on conditions where some of the speaker's and hearer's assumptions about the context differ. Two interlocutors may, for instance, disagree in what is relevant in the given situation without noticing this mismatch in their cognitive environments: the speaker may assume that 'spatial motion' is relevant in the given situation, and that 'intention' is not, while the hearer may think that the opposite is the case. The speaker then uses *be going to* in its conventional sense to communicate spatial motion. She thus only relies on the putative common ground knowledge that (i) the expression of motion is relevant in the given context, and (ii) that the hearer shares her convention which maps the meaning 'spatial motion' onto the form *be going to*. Whether or not she also considers the knowledge that spatial motion often implies intention to be common ground in that situation is not important, since, from her perspective, this knowledge would not lead to the inference of additional relevant information. The hearer, on the other side, just as in the metaphor-based scenario, will think that the speaker has used *be going to* metaphorically: he will reason that information about spatial motion is not relevant in the given context, but that the relevant information that intention is implied can be inferred. This scenario is illustrated in Fig. 3.6(a), and the speaker's and the hearer's reasoning is detailed in (6).

(6) Speaker's reasoning:

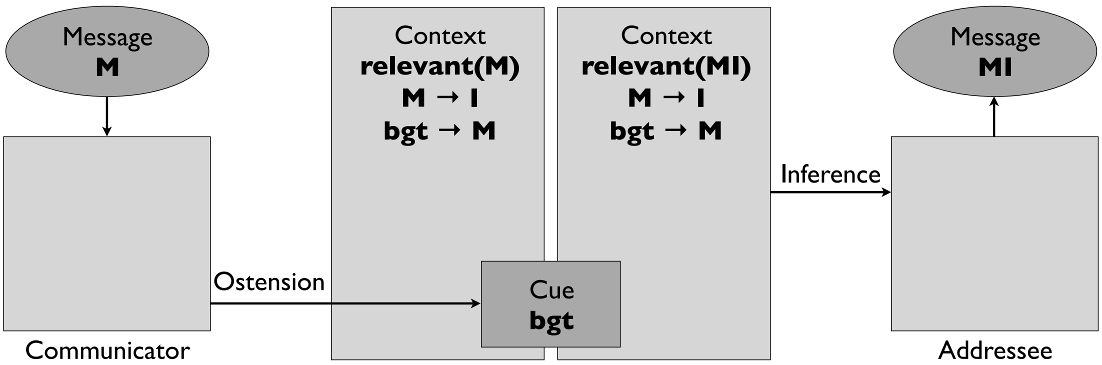
a. I want to express 'spatial motion'.
b. I have a construction for the expression of 'spatial motion' in my linguistic code, and the hearer shares this convention.
c. 'Spatial motion' is relevant in the given context.
d. ('Spatial motion' is associated with 'intention', but this does not matter because of (e).)
e. ('Intention' is irrelevant in the given context.)
f. Because we share common ground, the hearer will be aware of (b)–(e) and realise that I am aware of it too.
g. Because of (f), I can use the construction to communicate 'spatial motion'.

The hearer's reasoning is the same as in (5h)–(5n) above:

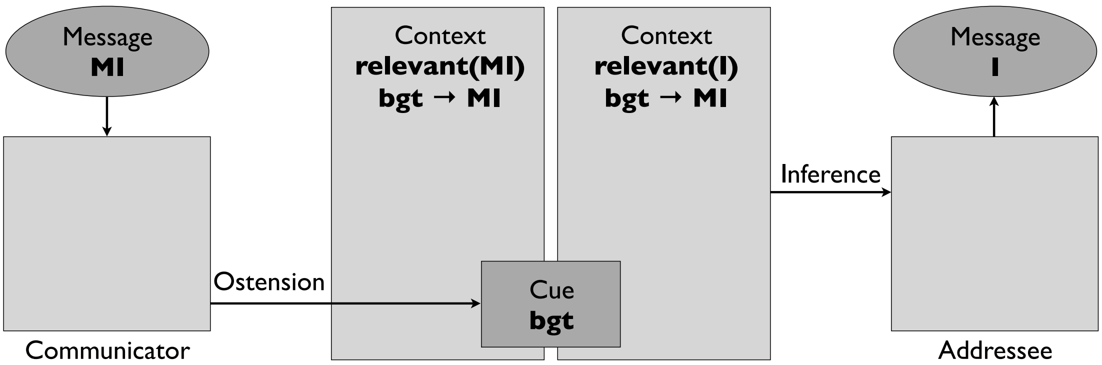
h. The speaker has expressed 'spatial motion'.



(a)



(b)



(c)

Figure 3.6: Reanalysis-based scenarios of grammaticalisation of *be going to*. Legend: *I* = 'intention', *M* = 'spatial motion', *bgt* = *be going to*. (a) The reanalysis-based scenario in one step. (b)–(c) The reanalysis-based scenario in two steps.

i. ‘Spatial motion’ is not relevant in the given context.
j. ‘Spatial motion’ often implies ‘intention’.
k. ‘Intention’ would be relevant in the given context.
l. I must assume that the speaker is cooperative.
m. I must also assume that the speaker is aware that I know (h)–(l), and that I know of his being aware of it.
n. From (h)–(m), I conclude that the speaker intends to convey ‘intention’.

Like the metaphor-based scenario, the reanalysis-based scenario too can be interpreted as a matter of two rather than just one usage event initiating change. After the interaction described above, the speaker will only further entrench the convention for *be going to* she already possessed before the usage event. The hearer, however, will remember that *be going to* has been used to convey intention, and thus entrench a new form-meaning association. How one can get to this result via *two* usage events with contextual mismatches is illustrated in Figs. 3.6(b) and 3.6(c). In a first communicative situation, the speaker does not think that intention is relevant but the hearer does. Both agree that information about spatial motion is relevant. In this situation, when the speaker uses *be going to* in its conventional sense to express spatial motion, the hearer will additionally infer ‘intention’ as relevant information that is being conveyed. He will thus entrench the usage of *be going to* to communicate a combination of spatial motion and intention. This usage event is represented in Fig. 3.6(b). In a later usage event, shown in Fig. 3.6(c), the hearer may himself have become the speaker. He will use the new convention for *be going to* to express spatial motion in combination with intention. The new hearer, however, may deem the aspect of spatial motion irrelevant in the given context, and thus only memorise the use of *be going to* to communicate intention. Again, both the one-step and the two-step scenario are based on the same cognitive mechanisms and can—under specific conditions—eventually lead to the same result. In the two-step scenario, the hearer mistakenly assumes underspecified use in the first usage event, and mistakenly assumes overspecified use in the second. In the one-step scenario, these two usages are combined in one and the same interaction.

A contextual mismatch also occurs if the hearer does not possess the knowledge of a linguistic convention that the speaker has, but the speaker assumes that the respective linguistic convention is common ground. Even in such a case, communication can theoretically be achieved if the hearer is able to infer the meaning from the context anyway. If the speaker uses the convention in an under- or overspecified way, the hearer will end up with that new form-meaning association but not with the original convention. And, of course, any combination of such a mismatch in the knowledge of conventions and a mismatch in the interlocutors' judgement of what is relevant can theoretically occur. Cases where the hearer lacks assumed knowledge of linguistic conventions typically occur in *language acquisition*. In these cases, naturally, there is no layering (the polysemous co-existence of an old and a new mapping) in the linguistic knowledge of the hearer after the usage event.

3.3.2.3 A unified approach to grammaticalisation

The crucial point to note is that the hearer does exactly the same reasoning in both what I have called the metaphor-based scenario and the reanalysis-based scenario. One could say that from the hearer's perspective both scenarios are metaphor-based, only the speaker uses *be going to* innovatively in the metaphor-based scenario but conventionally in the reanalysis-based scenario. Once more (see also 2.3.3.2), the comprehension-driven model of the cultural evolution therefore appears as a special case of the production-driven model: as soon as one accepts the reanalysis-based scenario, one also accepts the possibility for the metaphor-based scenario to happen because a hearer will only infer what a speaker could, at least theoretically, invite him to infer.

The only difference between the two approaches is whether, in the given situation, speaker and hearer make the same assumptions about what is common ground. At first glance, one might conclude that scenarios that emphasise contextual mismatch do not require the capacity to establish and recognise common ground. This conclusion, however, is not warranted because even in the reanalysis-based scenarios, the speaker and the hearer invite and draw their inference on the basis of what they *assume* to be common ground. The recognition of common ground, and especially of what is relevant in a given situation, is thus pivotal to both accounts of grammaticalisation. Neither of the two scenarios can therefore be elevated to the status of the "right" scenario: they are merely different situation-dependent combinations of the same set of cognitive processes. In this sense, the analysis presented here represents a unified approach to grammaticalisation.

In summary, I have sketched an approach to the emergence of grammar in which the type of semantic change found in grammaticalisation is introduced by over- and underspecified use. Grammatical constructions emerge like all other symbols too as the result of the same usage-induced semantic change that is also responsible for the emergence of symbolism. I have made this visible by abstracting away from linguistic description to the psychological underpinnings of the processes of syntacticisation and grammaticalisation. I have shown how both processes can be analysed as the product of iterated ostensive-inferential communication. The sketched approach to the emergence of grammar is thus consistent with the initial hypothesis, namely that (i) the emergence of schematic constructions and (ii) the process through which constructions become more functional originate from the same cognitive mechanisms as symbolism.

3.4 The protolanguage debate from a usage-based perspective

In recent years, the notion of "protolanguage" has become one of the most hotly debated issues in evolutionary linguistics. The two camps in the dispute are the proponents of a synthetic protolanguage (e.g. Bickerton 1990, 1995, 2003; Jackendoff 1999; Hurford 2003; Gil 2006; Tallerman 2007) and those who advocate a holistic protolanguage (e.g. Carstairs-McCarthy 1999, 2000; Wray 2000; Kirby 2000; Arbib 2005). Common to both approaches is that they distinguish two steps in the evolution of language: (i) the emergence of symbolism, and (ii) the emergence of grammar. The term "protolanguage" denotes the communication system that they assume was in place between the two steps, that is, after the emergence of symbolism but before the emergence of grammar. Protolanguage can thus be described as a system of symbols ("proto-symbols") without syntactic structure.

The protolanguage debate revolves around two questions. The first asks about the nature of protolanguage symbols. The two camps disagree on the phonological complexity of their forms, but above all on the semantic complexity of the meanings of these proto-symbols. The second question follows from the first: how do we get from protolanguage symbols to human language as we know it today? The answer to this second question is constrained by the answer given to the first. Depending on what the assumed nature of protolanguage symbols is, different processes must be postulated to explain the emergence of grammar on top of these symbols.

It seems appropriate to discuss this issue, which is at the center of so many discussions in the field, in the light of the usage-based perspective on the emergence of symbolism and grammar presented in this chapter. I will mainly address two questions. The first deals with the main point of disagreement between the two competing approaches: the nature of protolanguage symbols. What does the usage-based account developed in this chapter say about the semantic complexity of the first symbols? The second question concerns a more fundamental issue: the one assumption both sides of the dispute agree on—and which actually creates the ground for the whole debate—is that symbolism and grammar have emerged in two distinct evolutionary steps: symbolism first, grammar later. But what are the consequences for the protolanguage debate if, as I have argued in this chapter, this strict distinction between the emergence of symbolism and the emergence of grammatical structure cannot be maintained? These two questions will be addressed after a brief introduction to the two competing theories within the dispute, the synthetic and the holistic account of protolanguage.

3.4.1 *Synthetic vs. holistic protolanguage*

The main difference between the synthetic approach and the holistic approach to protolanguage is what they assume the meanings of protolanguage symbols to have been like. Those who advocate the synthetic view think that early symbols had simple, word-like meanings, that is, that they stood for individual concepts. The symbols of a synthetic protolanguage are thus very much like modern-day content words: they refer to concepts like, for instance, 'bear', 'fire', 'cave', 'axe' or 'hunt', 'eat', 'give'. The proponents of the holistic approach, on the other side, assume that protolanguage symbols had complex, sentence-like meanings, that is, that they stood for whole propositions. The symbols of holistic protolanguage thus contain meanings which are more similar to modern-day sentences rather than to individual words, like, for instance, 'give that to me' or 'stay away from my territory'. The protolanguage dispute is therefore, at its core, a dispute over whether protolanguage symbols had simple, word-like meanings representing general concepts, or whether they had complex, sentence-like meanings representing specific propositions.

The remaining differences between the synthetic and the holistic approach to protolanguage ultimately result from the opposing views that the two accounts have on the nature of protolanguage symbols. Because they make different assumptions about the nature of protolanguage, the two competing approaches also have to choose different pathways to explain how protolanguage could have turned into language as we know it today. Smith (2008:101) points out that their respective challenges can roughly be summarised as follows: the synthetic approach faces the problem of composing word-like units into sentences whereas the holistic approach needs to break sentence-like units apart into words.

The symbols of a synthetic protolanguage can be used in two ways: either in an isolated way, that is, to draw the addressee's attention to a particular object in the environment, a situation or event, or they can be *concatenated* in an *ad hoc* fashion to express more complex meanings—hence the term "synthetic." However, such ad hoc concatenation must not be confused with syntactic structure. Synthetic protolanguage is "a lexicon without syntax" (Bickerton 1995:51) and concatenation, as described by Bickerton, is merely a "slow, clumsy, ad hoc stringing together of symbols" (Bickerton 1995:65). Hurford (2003:53) also characterises synthetic protolanguage as a "a crude pidgin-like form of communication with no function words or grammatical morphemes."

Proponents of a synthetic protolanguage have come up with different accounts of the emergence of grammatical structure, at least partly influenced by whether they assume linguistic innateness or not. Some (e.g. Hurford 2003; Tomasello 2003b) invoke cultural processes, particularly grammaticalisation, to explain how protolanguage could have brought about grammar. For these researchers, the concatenation of individual symbols really is the crucial momentum that initiates the emergence of syntactic structure. Not so for Bickerton, who was among the first to propose a synthetic protolanguage on the basis of his study of pidgins (e.g. Bickerton 1984, 1990). Because he believes that syntactic structure is determined genetically, he must assume an additional biological step. For Bickerton, the fact that protolanguage users began to concatenate symbols to express more complex meaning is therefore not directly involved in the emergence of syntax.

The process invoked to account for the emergence of syntactic structure by those who advocate a holistic protolanguage is analytical *segmentation*. Grammatical patterns and word-like symbols emerge because users begin to generalise when, by chance, elements of form and elements of meaning coincide in two or more symbols. Wray (2000) provides the example of a hypothetical protolanguage that contains the symbols *tebima*, meaning 'give that to her', and *kumapi*, meaning 'share this with her'. She then assumes that users would detect the chance co-occurrence of the element of form *ma* and the element of meaning '(to/with) her'.

They would infer a new symbol *ma* which denotes a female beneficiary. Likewise, they would also create a grammatical construction *tebi*-X, meaning 'give that to X', and a construction *ku*-X-*pi*, meaning 'share this with X'. Wray's example is summarised below: (7) gives the protolanguage situation, and (8) renders the result of the segmentation process.

(7) a. tebima
'give that to her'
b. kumapi
'share this with her'

(8) a. tebi-X
'give that to X'
b. ku-X-pi
'share this with X'
c. ma
'(to/with) her'

Both camps, those who favour the synthetic approach and those who favour the holistic approach, claim that traces of protolanguage can still be found in present-day language. Wray (2000:286f.) refers to the fact that language contains "prestored multiword units for quick retrieval, with no need to apply grammar rules" and concludes from this that such formulaic sequences, as she calls them, show that present-day language still functions in part holistically. Bickerton, on the other side, not only thinks that protolanguage can be observed in pidgins (Bickerton 1984, 1990) but also points to the relationship between synthetic protolanguage and early child language:

> Until quite recently, it was generally assumed that ontogeny and phylogeny, though far from indissolubly wedded, were at least alike to this extent: the earliest units of pre-human utterances were pretty similar the earliest units of contemporary infants. That is to say, they were basically single units with ostensibly definable referents, perhaps somewhat broader in meaning than the units of an adult vocabulary. (Bickerton 2003:84)

The main characteristics of the synthetic and the analytic approach to protolanguage are summarised in Table 3.2. We can now discuss how the usage-based

	synthetic protolanguage	**holistic protolanguage**
meanings	word-like, simple, general, individual concepts	sentence-like, complex, specific, whole propositions
usage	concatenated or isolated	isolated
present-day equivalents	pidgins, child language	formulaic language
emergence of grammar	grammaticalisation or biological mutation	segmentation

Table 3.2: The main characteristics of the two competing approaches to protolangauge.

approach presented in this chapter fits in, and what consequences its implications bear for the protolanguage debate in general.

My assessment of the debate is that not enough attention has been given to the distinction between signal meaning and speaker meaning. Remember that speaker meaning refers to the information that the speaker actually intends to communicate in a particular communicative act. Signal meaning, on the other side, denotes the meaning that is conventionally associated with the produced linguistic form in the user's linguistic knowledge, the meaning that is encoded in a linguistic signal. The two approaches to protolanguage mainly differ in the extent to which they assume that protolanguage signal meaning was identical with protolanguage speaker meaning. Synthetic protolanguage encodes little and implies much, while holistic protolanguage encodes much and implies little. This is expressed in examples (9) and (10), which show how the same speaker meaning would be communicated by means of a synthetic protolanguage (9) and a holistic protolanguage (10) respectively.

(9) Synthetic protolanguage:
  a. Signal: *ugluk!*
  b. Signal meaning: 'apple!'
  c. Speaker meaning: 'Give me that apple!'

(10) Holistic protolanguage:

a. Signal: *ugluk!*
b. Signal meaning: 'Give me that apple!'
c. Speaker meaning: 'Give me that apple!'

Confusion may arise when proponents of the two competing theories do not make clear which sort of meaning they are talking about at a specific point. It is important to note that both approaches make the same assumptions about the complexity of protolanguage speaker meanings, they only differ in what they assume about the complexity of protolanguage signal meanings. Holistic approaches are maybe a bit more prone to confusing the two types of meaning because there is much less of a difference between the speaker meaning and the signal meaning in a holistic protolanguage than between the speaker meaning and the signal meaning in a synthetic protolanguage.[17] The dispute, however, evolves not around what protolanguage users were able to communicate but about what protolanguage symbols *conventionally* stood for.

3.4.2 The complexity of proto-meanings

One thing that a usage-based approach like the one developed in this chapter shows is that conventionality is a matter of degree. A convention emerges when the usage of a cue is remembered by two individuals and when these two individuals recognise that memory as common ground in later situations. For a meaning to be conventional, it thus has to have been communicated by virtue of the respective form at least once, and the memory of this usage event needs to be shared by at least two individuals. I doubt, however, that this is the sort of linguistic convention protolanguage theorists talk of. But if conventionalisation is a matter of degree, where do we draw the dividing line? After how many instances of re-use does a form-meaning association become part of the assumed protolanguage? Or how many individuals have to recognise a form-meaning association as common ground for that association to count as an element of the respective protolanguage?

[17] It needs to be noted, however, that contrary to the impression that may be created by the example, even the signal meanings of a holistic protolanguage can never fully specify the speaker meaning they communicate: at least reference resolution (symbol grounding) in the given context needs to be performed by the interlocutors. In the example, the mentioned apple needs to be identified in the specific physical environment of the interlocutors.

From a usage-based perspective, the question about the nature of protolanguage symbols must be reformulated along two of its aspects. First, conventional meanings are "fossilised" speaker meanings: conventions arise when usage situations are memorised. The first conventions must have arisen from usage situations where no convention was in place yet, that is, where the speaker meaning was inferred from the context without the support of a conventional meaning associated with the produced cue. One question that one can ask is thus what type of meanings could have been communicated when no conventions had been established yet, that is to say, how complex the information is that two individuals can exchange when they share no form-meaning associations. Second, in order to be considered part of some language by a hypothetical observer, form-meaning associations need to be used with a certain, albeit indeterminable, frequency and be shared by a certain number of people. A second question to be asked is thus how complex the meaning in a form-meaning association must or can be in order for the association to "catch on" and become what an observer might perceive as a relatively established part of a community's linguistic knowledge.

What can be said about the nature of speaker meanings that are conveyed in ostensive-inferential communication without the help of extant conventions? From a usage-based perspective, it is evident that first and foremost they must fulfil two conditions: (i) they must be *relevant* in the given situation, and (ii) they must be *inferable* from the context. The latter in turn implies that, in the given situation, there must be a cue available which allows the respective speaker meaning to be inferred. The complexity of ostensive-inferentially communicated meanings is thus situation-specific. As long as they are relevant and inferable, they can be simple or complex. However, what *can* be said about such meanings is that they are always fully *grounded*: these speaker meanings are always situated in the context of use. They do not convey concepts and categories but draw the addressee's attention to particular objects or events. A freshly entrenched form-meaning association will thus contain a very situated meaning, specific to the particular usage event of which it is a memory.

Whether or not such a new form-meaning association will "catch on" in a community depends on its re-usability. If, for instance, when a particular cue was used for the first time, it referred to a particular individual, and it is later re-used as a proper name for that individual, a simple and very specific meaning has become a convention. On the other side, if form-meaning associations are later re-used in an overspecified way, they become less specific. The first times, *ugluk* is used to express 'give me that apple', it is used in a situation where it refers to a specific apple, a specific "me," and so on. Later, it may be re-used in an overspecified way—if the memorised usage event exhibits a sufficiently salient similarity to the respective situation. The specifics of the original situation may then be ignored, and *ugluk* may come to mean 'give me that apple', independently of the context. Through further overspecified use, it may again change into 'give me that' or into 'apple', depending on what is relevant in the respective usage events.

The bottom line is therefore this. From a usage-based perspective, "proto-meanings" could have been either simple or complex—there is no reason to assume that only the one or the other could have existed at any one time. The condition a meaning has to fulfil in order to become conventionalised is that it is relevant and inferable in an initial usage situation, and that it is re-usable (and re-used) sufficiently often in later usage situations. Any form-meaning association initially represents a situation-specific memory and only gradually becomes more general through overspecified use, for which it needs to bear recognisable similarity to later usage situations. But after how many re-usages a particular form-meaning association is considered a convention is ultimately an arbitrary decision.

3.4.3 The protolanguage assumption

While no sides can be taken from a usage-based perspective regarding the complexity of protolanguage meanings—other than that both simple and complex meanings can be present at the same time—in terms of the processes employed, the account presented in this chapter seems to be closer to the view held by proponents of a synthetic protolanguage: it also emphasises that symbols can be concatenated. There are, however, some differences. One is with the Bickertonian account: while Bickerton recognises that the concatenation of protolanguage symbols must have been possible for protolanguage to be useful, he does not consider that concatenations can be conventionalised too. He assumes that syntactic structure is innate and must thus have been the product of a later biological mutation. In contrast, the ostensive-inferential model of communication suggests that concatenation itself serves as a cue to trigger the inference of meaning from context too, and can thus also become associated with that meaning like any other cue. The other point which thus deserves emphasising is that since the use of concatenations as cues is not different from the use of other cues, and since the conventionalisation of concatenations and the meanings they have been used to convey is not different from the conventionalisation of other cues and the meanings they have been used to convey, there is no reason to speak of a two-step or two-process scenario of language evolution.

The assumption of a protolanguage stage where all forms were associated with either simple or complex meanings does not make sense from a usage-based perspective of language evolution, and neither does one where no concatenations have become conventionalised. At any one stage, an individual's linguistic knowledge can contain simple and complex, specific and general, and more or less grammatical constructions, with different degrees of entrenchment. To assume an intermediate step in language evolution, protolanguage, appears futile if one considers the various possible scenarios of ostensive-inferential communication. Coming from a different angle, Smith (2006c:322) makes a similar call for a "more pluralistic conception of the evolution of language" than is assumed by the two-step protolanguage scenarios.

From the usage-based perspective introduced in this chapter, the protolanguage debate is thus flawed for three reasons. First, it fails to acknowledge that conventionality is a matter of degree. Second, it assumes an implausibly uniform intermediate stage in language evolution, "protolanguage." And third, it takes for granted the view that the mechanisms at the origin of grammatical structure are different from those eliciting symbolism.

3.5 Conclusion

This chapter has dealt with the emergence puzzle: language has arisen from no language. By identifying linguistic communication as a special case of ostensive-inferential communication, I have been able to link the "state before" with the "state after." The continuous element which is present in both is ostensive-inferential communication, which is possible with or without the application of conventionalised linguistic cues. This facilitates significantly the task of accounting for the evolutionary gap between no language and language: in essence, it is a matter of how linguistic conventions can emerge from ostensive-inferential communication.

I have suggested that forms become associated with meanings when particular usages of cues employed in ostensive-inferential communication are memorised.

Such form-meaning associations are conventionalised if they become part of at least two individuals' common ground and are used themselves as contextual information to support the inference of meanings in later usage situations. Furthermore, through frequent use, form-meaning associations are entrenched, turn into psychological units and can thus be accessed more automatically with much less (or no) situation-specific reasoning being necessary. The core property involved is pragmatic plasticity: cues can under- and/or overspecify the meanings they communicate in particular usage situations. The conventionalisation of under- and overspecified usages of extant form-meaning associations leads to semantic change: narrowing or broadening respectively. This is also how non-arbitrary form-meaning associations become arbitrary, and thus how symbolism emerges. I have suggested that this can also include syntactic constructions: that they are symbolic too, and therefore also emerge like any other symbols too. Syntacticisation can be conceived as a normal process of conventionalisation of the concatenation of symbols—where concatenation is itself employed as a cue that conveys meaning. Grammaticalisation can then be described as a special type of semantic change by means of which conventional meanings become more general until they can be viewed as "functional." Such semantic change is also usage-induced: it too comes about through under- and overspecified use and subsequent conventionalisation.

At the beginning of this chapter, I hypothesised that the common assumption made in evolutionary linguistics that symbolism and grammar have different origins may not be warranted. This hypothesis would suggest that not only can both the emergence of symbols and the emergence of grammatical structure be explained as products of cultural evolution but they are also not two distinct phenomena: both symbolism and grammar would emerge from ostensive-inferential communication. The cognitive capacities required for this to happen are (i) the ability to recognise common ground, and in particular (ii) the ability to recognise what is relevant in a particular interactive situation, as well as (iii) the ability to remember shared experience and to later refer to it. The approach sketched is compatible with the account for the emergence of symbolism given by Tomasello (1999, 2003a,b). It also agrees with Tomasello's view that grammar has emerged through processes of syntacticisation and grammaticalisation. However, I take this account a step further by arguing that these two processes, syntacticisation and grammaticalisation, could well be based on the same cognitive mechanisms that also lead to the emergence of symbolism.

In this chapter, I have refined the production-driven model of the cultural evolution of language introduced in the previous chapter. In particular, I have further specified how innovation comes about as the result of use. The presented account has the characteristics that Kemmer and Barlow (2000) list for usage-based models of language: (i) there is an intimate relation between linguistic structures and instances of use; (ii) frequency plays an important role; (iii) comprehension and production are integral, rather than peripheral, to the linguistic system; (iv) language learning focuses on experience; (v) linguistic representations are emergent, rather than stored as fixed entities; (vi) the observation of usage is important for the construction of the model; (vii) there is an intimate relation between usage, synchronic variation, and diachronic change; (viii) the linguistic system is interconnected with non-linguistic cognitive systems; and (ix) context plays a crucial role in the operation of the linguistic system.

One consequence of this usage-based perspective on language evolution is that it leaves little reason to assume a distinct intermediate step between no language and language, a protolanguage. Rather, language continually emerges and changes. In order to study how, in the course of this process, language comes to exhibit the appearance of design for communication, I will devise computer simulations that are based on the usage-based model introduced here. In the next chapter, I will therefore propose a computational implementation of this conceptual framework.

CHAPTER 4

Computational implementation

The study of language evolution has often been supported by computational models and simulations. Computer simulations have been used to enhance the study of phenomena as diverse as the evolution of sound systems (De Boer 2002, 2005), the evolution of linguistic structure (Brighton 2002; Batali 2002; Christiansen et al. 2002; Kirby 2002a; Kirby and Hurford 2002; Smith et al. 2003), the evolution of linguistic diversity (Livingstone 2002; Solan et al. 2005), how language adapts to be learnable (Brighton et al. 2005), symbol grounding and the creation of meaning (Cangelosi et al. 2002; Steels 2002; Steels and Kaplan 2002; Smith 2005a), and the co-evolution of language and the language faculty (Briscoe 2002a, 2005). But why does it make sense to complement conceptual models of language evolution with computer simulations?

Computational modelling has proved able to serve three main purposes: (i) theory refinement, (ii) hypothesis testing, and (iii) laboratory-like experimenting (cf. Cangelosi and Parisi 2002a:8–12 for a similar analysis). First, computational models can lead to theory refinement because they cannot but be fully specified. A computational model forces its developer to think of every case or state of affairs that can possibly occur, and to specify what their theory has to say about that case. To transform a developed theory into a computational model thus helps researchers to discover and fill holes and inconsistencies in their conceptual thinking. Second, a computational model allows for the testing of predictions made by a theory. If they are formulated specifically enough, such hypotheses can be put to test by running computer simulations. At best, if the simulation behaves in the predicted way, a hypothesis can be corroborated; at worst, it must be refuted, or previously unconsidered factors must be taken into account. The latter case points to the third function of computational models, namely to provide a laboratory-like experimental environment. In the process of attempting to find a configuration in which the simulations yield the expected result by a somewhat directed technique of trial-and-error, the experimenter may become aware of previously overlooked parameters that have an impact on the behaviour of the modelled system.

Because they provide the opportunity for hypothesis testing and laboratory-like experimenting, computational models are particularly helpful in cases where (i) the phenomenon to be investigated cannot be observed directly, and/or (ii) where complex adaptive systems are involved in the said phenomenon. Language evolution exhibits both these conditions. First, the origins of language are simply too remote in the past for any direct observation of the process to be available. And even though some present-day processes of language use and change may provide insights into how language could have emerged in the first place, the actual emergence of linguistic communication from a fully non-linguistic environment cannot be observed anymore. Computational models and simulations thus provide a way of dealing with the notorious data problem of evolutionary linguistics (cf. Cangelosi and Parisi 2002a:4). Second, language—particularly in its diachronic aspects—has frequently been described as a complex adaptive system (e.g. Gell-Mann 1992; Kirby 1999, 2002b; Hashimoto 2002; Holland 2005). Complex adaptive systems involve multiple local interactions which, in their sum, lead to emergent structure on a global level. They change over time, and earlier states influence later states. These properties—the interaction of multiple local processes and the accumulation of effects over time—make it difficult to predict the behaviour of complex adaptive systems. The behaviour of an individual termite, for instance, does not allow one to infer the shape of the mound that results from the continuous interactions of a whole colony of these insects. Similar examples can be found in the flocking of birds, the behaviour of the stock market or the occurrence of traffic jams. Likewise, the shape of a language is hard to predict from the specifics of individual human communicative interactions, and this is where computer simulations can be of help.

Computational models necessarily include simplifications and idealisations. If the effect of a particular phenomenon is to be studied with the help of a simulation, that phenomenon has to be modelled in isolation: other phenomena must be simplified and abstracted away from so that they cannot interfere with the outcome of the simulation and obscure the conclusions that can be drawn about the effect the phenomenon in question has. Only if other possible causes are excluded from a simulation can one be sure that the observed effect really is caused by that phenomenon. Suppose we hypothesise that a phenomenon A can cause an effect X. To test this hypothesis by means of simulations, we have to exclude any phenomenon B that could also cause X. Only then can we be sure that if X emerges in the course of the simulation, it was caused by A. Kirby (1999:44, emphasis in the original) explains this reason for why computational models necessarily include simplifications and idealisations with an example:

> The purpose of the simulation is not to be a complete analogue of the real world. Rather, it should be a reification of a theory. It should involve all the idealizations that a model of that theory would involve and only those idealizations. If we were to build a simulation of some theory of the flocking of birds, let us say, and we built in a detailed description of wind direction which the theory did not mention, then the results of the simulation tell us nothing about the validity of our original theory. Of course, the process of building and testing the simulation might lead us to conclude that the original theory does not work *without* taking into account wind direction, but this simply serves to underscore the importance of simulation.

As I have pointed out before, most computational models in the field have been designed to study the effect of *imperfect cultural transmission (learning)* on the evolution of language. Consequently, *communication (use)* has been modelled in an idealised and simplified way. However, because they abstract away from the mechanisms of communication, these models depend on (i) presupposing the existence of symbolism and (ii) assuming an artificial process of random invention (Hurford 2002). The first condition renders these models unable to account for the emergence of symbolism from a non-symbolic state. The second condition forces them to include what can be seen as something of a deus-ex-machina solution to linguistic creativity: random invention is very rare in real language-use (see e.g. Trask 2000:369).

In this chapter, I devise a computational model that focuses on the effect of communication (use)—and specifically the fact that signals exhibit pragmatic plasticity—on the cultural evolution of language. I will propose ways in which such a model can (i) simulate the emergence of a symbolic communication system from an initial non-symbolic (and hence pre-linguistic) state and (ii) avoid having to include a process of random invention. The computational model to be introduced will be based on the conceptual considerations made in chapters 2 and 3, and it will be put to use to investigate questions relating to the design puzzle later in chapter 5. The remainder of this chapter falls into two parts. I will first describe the general architecture of the computational model as a model of iterated *symbolic* communication (section 4.1). I will then show how we can use exactly the same model—without any modifications—to simulate *non-symbolic* communication, and the emergence of symbolism (section 4.2).

4.1 Modelling iterated symbolic communication

I begin by describing the general architecture of a computational model that simulates the iterated use of symbolic conventions. I will thereby understand symbolic communication in the broad sense as any form of communication that makes use of *conventional* associations between forms and meanings, be they arbitrary (i.e. "symbolic" in the narrow sense) or not. Later, in section 4.2, I will show that the same model is also capable of simulating non-symbolic communication, and that symbolism does therefore not need to be presupposed. But for the moment, I will discuss the model's architecture under the assumption that the existence of conventional form-meaning associations is given. In particular, I will look at what aspects of cultural evolution the model simulates, and describe how it realises the two main process of cultural evolution: use and learning.

4.1.1 Simulating cultural evolution

In section 2.3.1, I explained that the cultural evolution of language can be modelled as a process of iterated use and learning: I-language is mapped to E-language through use and E-language is mapped back to I-language through learning (Kirby and Hurford 2002:123). This is the central idea of the iterated learning model (ILM) in the broad sense. I have argued that there are two possible loci for the introduction of innovation in this model: innovation can come about in the course of use or in the course of learning. If, in a simulation, we want to study the effect of innovation in one of these processes, we have to model the other in an idealised form as a faithful process. It must be understood, however, that the two schemata that follow from this consideration, *faithful use / innovative learning* and *innovative use / faithful learning*, are merely designed to study the effects of each respective type of innovation in isolation. In reality, neither use nor learning can be assumed to be completely faithful. Most existing computational implementations of the ILM study the effect of innovation through imperfect learning and thus represent use as faithful. This setup maps the idea of the ILM onto Sperber's (1996) general model of cultural evolution, which also emphasises the role of imperfect cultural transmission. In contrast, the computational model to be introduced in this chapter is intended to study the effect of innovative use and will therefore represent learning as faithful. This setup, in turn, maps the idea of the ILM onto Tomasello's (1999) ratchet model of cultural evolution, which emphasises the relative fidelity of cultural transmission (cf. section 2.1.3.3).

The assumption of innovation in use and faithful learning allows for two further simplifications in the design of a computational model of linguistic cultural evolution. One is based on the fact that if learning is faithful, the role of social transmission in the process of cultural evolution is reduced to ensuring the preservation of innovations beyond a single individual's lifespan. An individual with a long (potentially unlimited) lifespan can theoretically accumulate the same innovations that a number of short-lived individuals do with the help of social transmission. This effect is visualised in Fig. 4.1: the innovations accumulated by the long-lived individual *X* alone are a projection of those accumulated by the short-lived individuals *A*–*D* together. Thus, if learning is assumed to be faithful, nothing can be gained from modelling social transmission—unless one wants to study the spread of an innovation in a population, which is not the purpose of this model (cf. the discussion of micro-dynamic change vs. macro-dynamic change presented in section 2.3 and the caption to Fig. 4.1). Because the present computational model assumes learning to be faithful, the cumulative effect of iterated social transmission can thus be simulated by projecting it onto an idealised individual with an unlimited lifespan as illustrated in Fig. 4.1.

A second simplification the assumption of innovative use and faithful learning makes possible is that only one of the two individuals involved in every act of language use, the speaker, needs to be modelled explicitly. Why is it that more than one individual is involved in the use of linguistic artefacts in the first place? For many cultural artefacts, only one individual is required to create a chain of use/learning-iterations as depicted in Fig. 4.1 above. An individual might for instance have knowledge of the use of a certain tool. To meet the requirements of a specific environment, he may apply this technique in a novel way (innovative

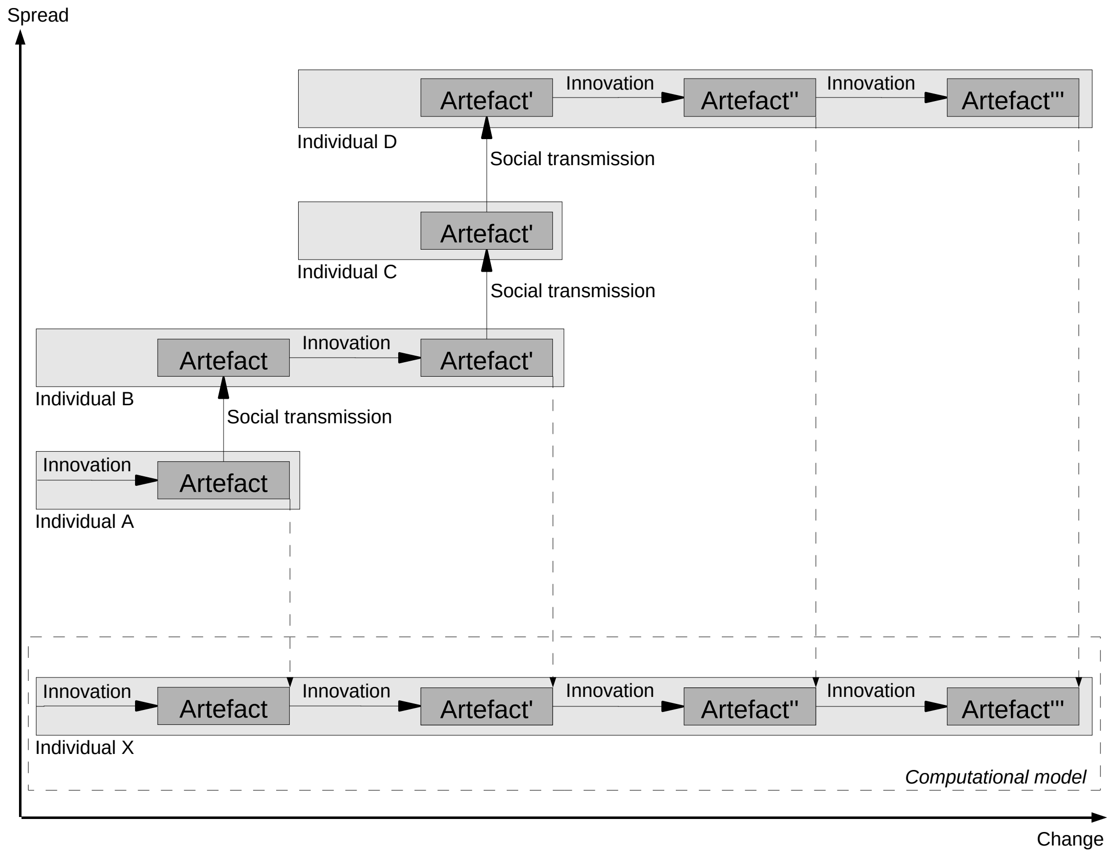


Figure 4.1: Simulating cumulative cultural evolution. The x-axis represents spread (macro-dynamic change), the y-axis change over time (micro-dynamic change). Faithful social transmission ensures the continuity of a chain of innovations beyond an individual's lifespan. Like this, a sequence of short-lived individuals can accumulate change that could only have been brought about by an extremely long-lived individual otherwise. In the present computational model, the cumulative effect of iterated faithful social transmission is thus simulated by projecting it onto an individual with an unlimited lifespan (as shown inside the dashed box).

use) and then memorise this modification of the original technique (faithful individual learning). In the same way, a language user can be seen to bring about a linguistic innovation. The difference however is that language is—as explained in section 3.1—a cultural artefact that is used to communicate some information to a second individual. Language use therefore always implies the presence of at least two individuals: a communicator (who, in our case, is also the innovator) and an addressee. In chapter 3, I have argued that linguistic communication is ostensive-inferential and therefore based on two interlocutors recognising (or at least assuming to recognise) common ground. I have also pointed out that the minimal setup for symbolic communication to evolve consists of two individuals repeatedly communicating with each other. This scenario is depicted

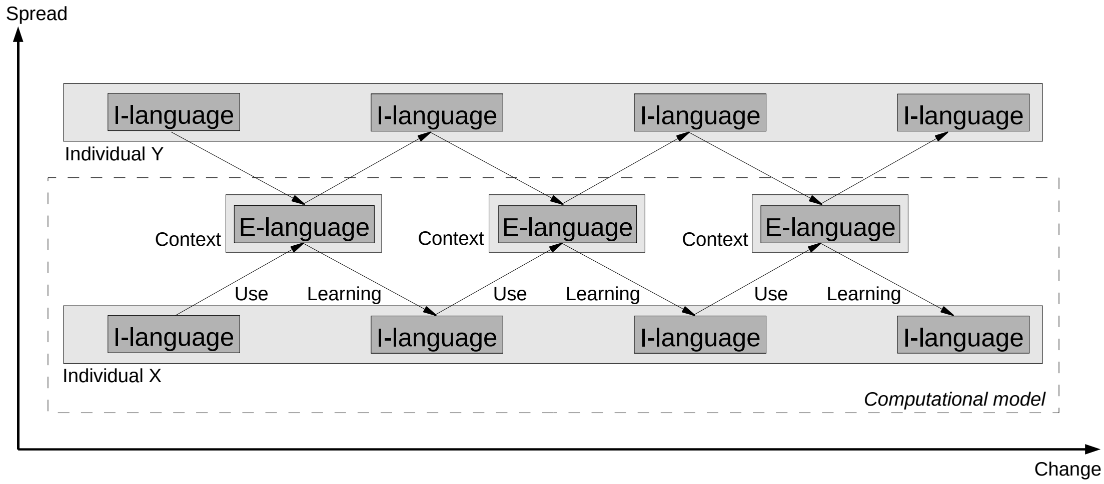


Figure 4.2: The accumulation of linguistic innovations through iterated language use and learning. The computational model simulates (i) the communicator iteratively producing E-language on the basis of his I-language and a given context (use) and updating his I-language on the basis of E-language (learning), and (ii) for each usage situation the contextual information the communicator assumes to be common ground shared with the addressee. The addressee is not represented explicitly but his presence is implied in each usage situation through the simulation of context.

in Fig. 4.2. Linguistic change comes about in both the communicator and the addressee through innovative use in context (production as well as comprehension) and subsequent updating of their I-language (i.e. individual learning in the communicator and social learning in the addressee). Fig. 4.2 shows that the computational model simulates (i) the communicator, and (ii) for each usage situation, the contextual information the communicator assumes to be common-ground knowledge shared with the addressee. The addressee is not modelled explicitly—note, however, that the presence of an addressee is implied in each usage situation by the fact that common-ground knowledge is available to the communicator.

The decision not to model the addressee explicitly has been made on the basis of the following considerations. The addressee will evolve an I-language that is different from the communicator's under one (or both) of two conditions. First, variation is introduced when the interlocutors' assumptions about what constitutes common ground differ in a usage situation. I have argued in section 2.3.3.2 that this scenario merely constitutes a complication of the default case, and that we should only resort to it once the explanatory power of that default case has been fully exploited. Second, even if we assume that the interlocutors always successfully recognise their common ground, they may end up with different I-languages if they start with different initial states. However, such a scenario runs counter to our aim to model the emergence of language from a non-linguistic state where neither of the interlocutors possesses any language yet. Under the given conditions, the I-language evolved by the addressee would thus always be identical to the communicator's, and we would not gain anything by modelling the addressee explicitly.

The general architecture of the computational model can now be described as follows. The model is designed to study the effect of innovative use on the cultural evolution of language. It simulates how the I-language of an individual evolves through iterated engagement of that individual (as speaker) in communicative acts. The implementation is composed of the following components:

- The individual and its I-language are modelled as an *agent* possessing a set of *form-meaning mappings*.[1]

A simulation run consists of a series of iterations.

- Each iteration represents a *communicative situation*, in which the agent is presented with (i) a context and (ii) a meaning he needs to communicate in this context.
  (Both context and meaning are generated randomly for each iteration.)

In each iteration, the agent then performs the following two acts:

- *Use (potentially innovative):* on the basis of (i) his I-language and (ii) the given context, the agent produces a signal for the meaning he needs to communicate.
- *Learning (faithful):* the agent then updates his I-language by adding or entrenching the association between the used signal and the communicated meaning.

I will now describe how the two processes of use and learning are modelled in more detail.

[1]For the moment, it will suffice to note that forms are represented as sets of phonological units and meanings are represented as sets of semantic units. Both phonological units and semantic units are indicated by upper-case characters in the model. As an example, $XY \rightarrow ABC$ stands for a symbolic association that maps a form *XY* onto a meaning *ABC*. A more extensive discussion of the representation of forms and meanings is provided in section 4.2.1 below.)

4.1.2 Use

The core aspect of the present computational model is that it allows use to be innovative. But what is innovative use and how can we simulate it? In order to answer this question, the fundamental distinction between signal meaning and speaker meaning has to be taken into account. Signal meaning is the meaning *conventionally* associated with a linguistic signal, and speaker meaning is the meaning a signal actually communicates in a specific context. Use is therefore conventional (i.e. non-innovative) if, in a specific usage situation, the speaker meaning is the same as the signal meaning. It follows that use is innovative (i.e. non-conventional) if the speaker meaning (at least partly) differs from the signal meaning, that is, if the actually communicated meaning deviates from the conventional meaning of the produced linguistic signal.

Innovative use is thus the product of *pragmatic plasticity*. In section 3.1, I have argued that language exhibits pragmatic plasticity because linguistic communication is *ostensive-inferential*: linguistic signals and their conventional meanings serve as clues (ostensive stimuli) that guide the inference of speaker meaning on the basis of common-ground knowledge recognised by speaker and hearer. I have also shown that a signal meaning can deviate from the actually communicated speaker meaning in two ways: it can underspecify the speaker meaning and/or it can overspecify the speaker meaning. The simulation of innovative use therefore requires the process of use to be modelled in a way that allows for underspecification and overspecification.

4.1.2.1 Simulating under- and overspecification

The signal *underspecifies* the speaker meaning if it constitutes only a part of the speaker meaning (and additional information needs to be inferred.) Sentence (1) provides an example of underspecification: (1a) represents the encoded signal meaning, (1b) the actually communicated speaker meaning. The relationship between John and his book is not made explicit in the signal meaning but needs to be inferred from the context: the expression *John's book* could refer not only to a book written by John but also to a book owned by John or even to a book about John.

(1) a. I enjoyed reading *John's book*.
 b. [I enjoyed reading *the book written by John*.]

In the computational model, I devise an abstract set-theoretic representation of meanings and forms: meanings are conceived as combinations (sets) of semantic units and forms as sets of phonological units. Both semantic and phonological units are symbolised by upper-case characters. The model does not specify what individual characters stand for: they are kept abstract and are only meant to represent the fact that there *are* different semantic units available to language users. But we can easily interpret them in terms of specific linguistic examples. The example of underspecification given in (1) can be expressed in the model's representation as shown in (2). The list in (2a) breaks the speaker meaning occurring in the example down into semantic units, and (2b) shows how the fact that the signal meaning underspecifies the speaker meaning can be modelled with this representation: signal meaning *JRB* constitutes only a part of the speaker meaning *JRWB* and the additional information *W* is inferred.

(2) a. *J* = *john(x)*
B = *book(y)*
R = *relation(x,y)*
W = *written_by(y,x)*
b. Signal meaning: *JRB*
Inferred meaning: *W*
Speaker meaning: *JRWB*

In contrast, the signal *overspecifies* the speaker meaning if only a part of it occurs in the speaker meaning (and the rest of it is ignored because it is irrelevant.) An example of this phenomenon can be found in the metaphorical use of the word *chameleon* in sentence (3). Again, (3a) illustrates the signal meaning and (3b) the actually communicated speaker meaning. Only one aspect of 'chameleon' contributes to the speaker meaning here: the fact that chameleons frequently change their appearance. Other aspects of it, for instance that chameleons are reptiles or that they have long tongues, are ignored because they are irrelevant in the given context.

(3) a. Sally is a chameleon.
b. [Sally frequently changes her appearance.]

Example (4) provides the respective abstract set-theoretic representation employed in the computational model. We can think of 'chameleon' as a complex semantic unit (C) that in turn consists of semantic units such as 'long tongue'

(T), 'reptile' (R), 'changes appearance' (A).[2] The signal meaning overspecifies the speaker meaning because some of its components (R, T, ...) are ignored and do not form part of the speaker meaning.

(4) a. S = *sally(x)*
C = *chameleon(x)*
A = *changes_appearance(x)*
R = *reptile(x)*
T = *has_long_tongue(x)*

b. Signal meaning: SC where $C = \{A, R, T, ...\}$
Ignored meaning: R, T, ...
Speaker meaning: SA

What the above analyses of underspecification and overspecification imply for the task of simulating innovative use is this: in order to use a signal innovatively, that is, in order to be able to under- and/or overspecify the intended speaker meaning, a speaker needs to know what constitutes

- *inferable meaning,*
  namely which components of the speaker meaning the hearer can infer from the context,

and she needs to know what constitutes

- *ignorable meaning,*
  namely what meaning will be ignored by the hearer because it is irrelevant in the present context.

The former needs to be known to determine if underspecification is possible, the latter to determine if overspecification is possible. If, for instance, a speaker knows that (i) he needs to convey a meaning AB, and he also knows that (ii) B can be inferred from the context (once A is given), and that (iii) C will be ignored in the present context because it is irrelevant, then he can infer that he can use signals with the signal meanings A, AB, AC or ABC to convey the speaker

[2]While this interpretation is a useful simplification for the sake of the computational model, it must be kept in mind that it is only an abstract way of representing the encyclopedic, network-like nature of meaning (e.g. Fillmore 1975; Haiman 1980a; Langacker 1987; Evans and Green 2006).

meaning in the current context. (Signal meaning *A* would underspecify the intended speaker meaning *AB*, signal meaning *ABC* would overspecify it, and signal meaning *AC* would both under- and overspecify the intended speaker meaning. Signal meaning *AB* would provide an exact specification.)

One can therefore conclude that in each iteration—remember that an iteration simulates an individual usage event—the agent needs to be provided with:

- a (randomly generated) *speaker meaning* that he needs to communicate
- a (randomly generated) *context*

From this information, the agent can then infer:

- what constitutes *inferable meaning*
- what constitutes *ignorable meaning*

I propose the following way of modelling this process of inferring what constitutes inferable and ignorable information in a given situation:

I have explained in section 3.1.3.1 of this thesis that one precondition of ostensive-inferential communication is for individuals to recognise what constitutes *relevant* information in the given situation. I have stated that relevant information is information that, if it was transferred from the communicator to the addressee, would elicit a reaction in the addressee, a perlocutionary effect, that could plausibly be intended by the communicator in this situation.

The fact that an individual knows what is relevant in a given situation can be modelled abstractly by, in each iteration, equipping the agent with a set of meanings that represents the set of all meanings that he recognises as being relevant in the situation simulated by that iteration. The agent may, for example, "know" that both (and only) the meanings *AB* and *BD* constitute relevant information in a given situation.

The agent's task is then to provide a *clue* that contains *sufficient information to identify unequivocally* the intended speaker meaning from among the meanings in this set.[3] Imagine again that that set consists of the meanings *AB* and *BD* and

[3]The fact that, in order for communication to succeed, the provided clue must contain sufficient information for the addressee to be able to identify the intended speaker meaning is captured in the so-called Q-principles formulated by Gricean pragmaticists (e.g. Levinson 2000; Horn 2004), which call on the speaker to "be sufficient." I provide a brief discussion of how pragmatic principles are realised in the computational model in section 4.1.2.3 below.

the speaker meaning that the agent needs to bring across is *AB*. What minimal information does the agent need to provide as a clue to unequivocally identify *AB* in the set {*AB, BD*}?[4]

- Obviously, *B* alone does not constitute sufficient information: it could point to either of the two meanings.
- In contrast, *A* alone provides sufficient information to identify the intended meaning unequivocally, as it does not occur in any other meaning in the set.
- Furthermore, once *A* has been specified, *B* is not needed anymore: it can be inferred once *A* is given. In the given context, *B* therefore constitutes *inferable* information.
- What if the clue that the agent provides is *AC*? This clue also provides sufficient information (*A*) to identify the intended meaning *AB* unequivocally. Its component *C* does not contribute to this process but it also does not hinder it as it does not point to any other relevant meaning either. In the given context, *C* constitutes *ignorable* information.
- The situation would be different if the agent provided *AD* as a clue. In this case, *A* would point to one of the possible meanings, *D* to another one. The intended meaning could not be identified unequivocally. As opposed to *C* above, *D* does therefore not constitute ignorable meaning in the given context: while, like *C*, it does not occur in the intended meaning, it does occur in another, not intended relevant meaning.

In summary, it can be stated that if there is (i) an intended speaker meaning and (ii) a set of relevant meanings in which the intended speaker meaning has to be identified unequivocally, we are able to infer deterministically what constitutes inferable and ignorable information:

- *inferable information*
  is any information that is contained in the intended meaning but is not apt (or not necessary) to identify that meaning unequivocally
- *ignorable information*
  is any information that does not occur in the intended meaning but also does not occur in any other relevant meaning.

[4]Note that it is always presupposed here that the addressee recognises the provided clue as a clue, that is, that he recognises the communicator's communicative intention and thus makes the presumption of relevance.

However, in our computational model, we can make a further simplification to the above scenario: instead of randomly generating a context which yields (deterministically) the required information (i.e. the information about what constitutes inferable and ignorable meaning), we can directly generate this information itself: in each iteration, we can provide the agent with a randomly generated set of inferable and ignorable semantic units. (With regard to the discussed example, this would mean that the agent would be provided with the knowledge that, in the given context, *B* constitutes inferable meaning and *C* constitutes ignorable meaning.)

We are now in a position where we can further detail the individual stages of an iteration in the simulation as shown in Figs. 4.3–4.6:

1. *I-language* (Fig. 4.3). The input to each iteration is the I-language that the agent has evolved in the process of the preceding iterations. In the example, the agent's I-language contains only a single construction, which maps a form *X* onto a meaning *AC*. The asterisk signifies that, from a theoretical perspective, the I-language too must be considered as part of the context of use even though it is implemented separately in the computational model (refer to section 3.1.4 for a detailed account of this theoretical point).

2. *Speaker meaning and context* (Fig. 4.4). At the beginning of each iteration, the model randomly generates (i) a speaker meaning for the agent to communicate (*AB* in the example), and two sets of semantic units that designate what constitutes (ii) inferable meaning and (iii) ignorable meaning.[5] In the figure, inferable meaning (*B*) is marked with a plus sign, ignorable meaning (*C*) with a minus sign.

3. *Use* (Fig. 4.5). The agent then uses this information to produce an appropriate signal.[6] In the example, the agent's use of the signal *X* is innovative: the conventional signal meaning (*AC*) is different from the communicated speaker meaning (*AB*).

4. *Learning* (Fig. 4.6). The agent then updates his I-language by adding a new construction that associates the used signal (*X*) to the communicated meaning (*AB*) and further entrenches the construction(s) he used to produce the

[5]Appendix A provides a detailled description of the algorithms that generate speaker meaning and context.
[6]Appendix A provides a detailled description of the employed signal production algorithm.

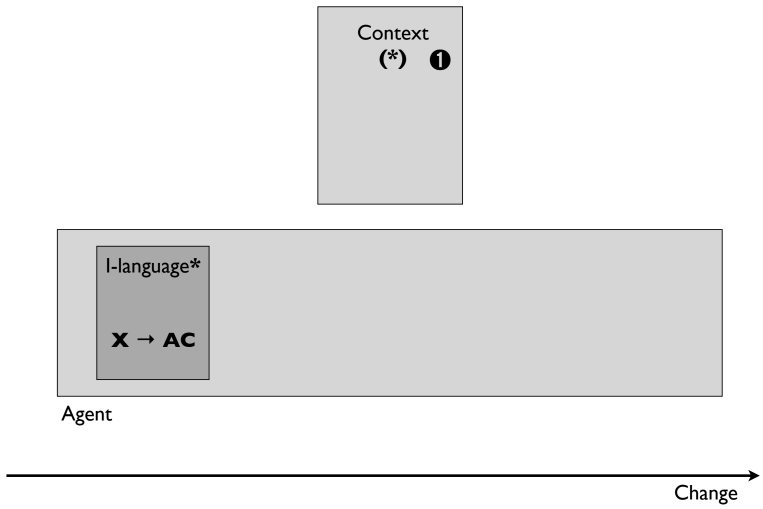


Figure 4.3: First stage of an example iteration (*I-language*). The output of the previous iteration (the agent's I-language) serves as input to the next iteration. The asterisk indicates that, theoretically, I-language too has to be conceived as part of the context of a usage event.

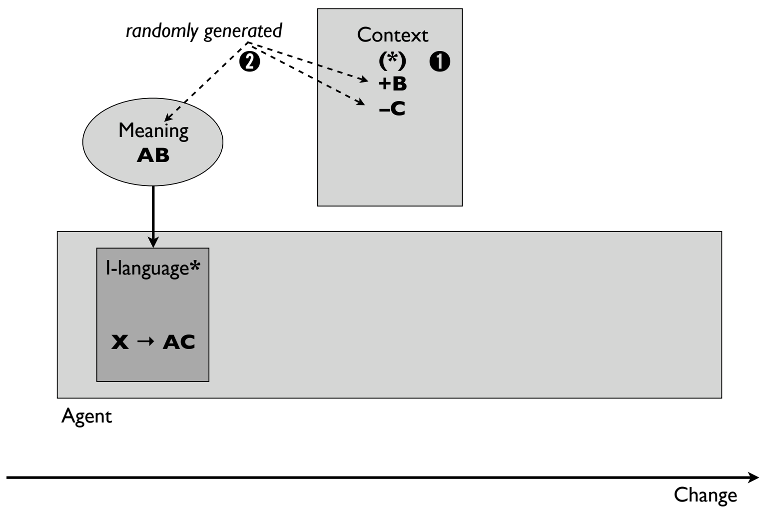


Figure 4.4: Second stage of an example iteration (*speaker meaning and context*). A speaker meaning, inferable meaning and ignorable meaning are randomly generated. (Inferable meaning is marked by a plus sign, ignorable meaning by a minus sign.)

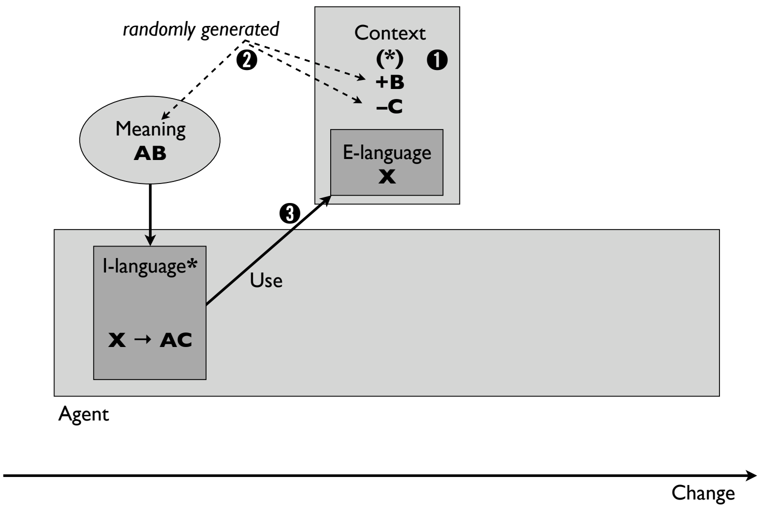


Figure 4.5: Third stage of an example iteration (*use*). The agent produces a signal that is appropriate to communicate the intended speaker meaning in the given context. The signal meaning can under- and/or overspecify the speaker meaning within the boundaries determined by the inferable meaning and the ignorable meaning.

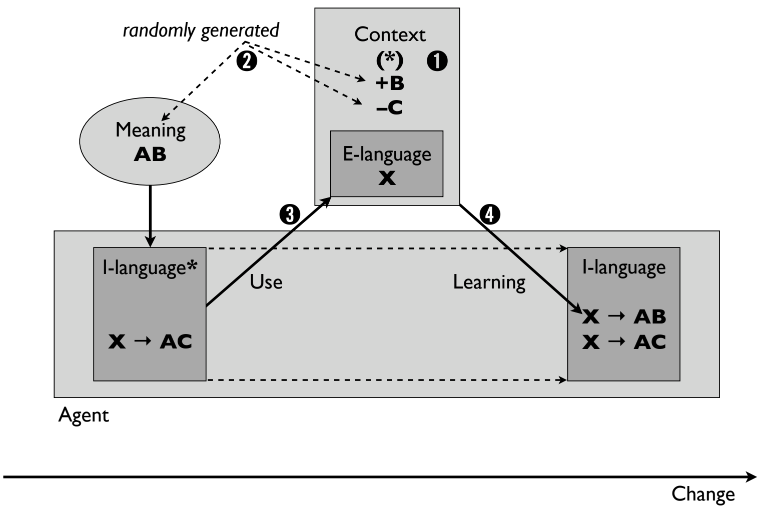


Figure 4.6: Fourth stage of an example iteration (*learning*). The agent updates his I-language by adding a new construction that associates the used signal with the communicated meaning.

signal ($X \rightarrow AC$). The process of learning will be explained in more detail in section 4.1.3 below.

4.1.2.2 *Forms of conventional and innovative use*

In order to illustrate the mechanisms of use employed in the computational model, I will now describe the basic forms of conventional and innovative use that can occur in the simulations. To this aim, I will present a series of examples and discuss for each of them (i) the conditions the agent meets in the respective interaction, (ii) how he decides on a signal on the basis of these conditions, and (iii) whether the resulting instance of use is conventional or innovative. The first example is given in (5) and illustrates the conventional use of a single symbol.

(5) *Conventional use of a single symbol*

Input:	I-language:	$X \rightarrow AB$
	Speaker meaning:	AB
	Inferable meaning:	–
	Ignorable meaning:	C
Calculations:	Possible signal meanings:	AB, ABC
	Expressible signal meaning:	AB
Output:	Signal:	X

This example shows a situation where the agent's I-language contains just one construction, mapping form X onto meaning AB. The agent needs to express speaker meaning AB, and he knows that no part of that meaning can be inferred and that C, if it was part of the signal meaning, would be ignored. In this context, both AB and ABC constitute signal meanings that contain sufficient information for the hearer to be able to infer the intended speaker meaning. However, only signal meaning AB can also be expressed with the help of the agent's I-language. The agent thus produces a signal X to communicate the intended speaker meaning AB. Because signal meaning and speaker meaning are equal, this example constitutes a case of conventional use.

The model also allows for symbols to be combined. This is shown in example (6):

(6) *Combination of symbols*

Input:	I-language:	$X \rightarrow A$
		$Y \rightarrow B$
	Speaker meaning:	AB
	Inferable meaning:	–
	Ignorable meaning:	C
Calculations:	Possible signal meanings:	AB, ABC
	Expressible signal meaning:	AB
Output:	Signal:	XY

The situation depicted in (6) is the same as in example (5) except for the fact that the agent's I-language this time makes it necessary for two symbols to be combined to express the required signal meaning. It is debatable whether this example should be seen as a case of conventional or innovative use. What speaks for a classification of it as conventional use is the fact that, like in the previous example, signal meaning and speaker meaning are the same. What speaks against it, however, is that the association between the used signal (*XY*) and the communicated meaning (*AB*) is not part of the speaker's I-language before the usage event: it is only added in the process of learning that follows the production of the signal. The two symbols are obviously being combined for the first time, which may itself be viewed as an instance of innovative use.

Example (7), in contrast, shows a clear case of innovative use: one in which underspecification is employed.

(7) *Underspecification*

Input:	I-language:	$X \rightarrow A$
	Speaker meaning:	AB
	Inferable meaning:	B
	Ignorable meaning:	–
Calculations:	Possible signal meanings:	A, AB
	Expressible signal meaning:	A
Output:	Signal:	X

This example shows a situation where the agent needs to convey a speaker meaning *AB* in a context where *B* can be inferred once *A* is given. A sufficient signal meaning must therefore at least express *A* but can optionally also specify *B*. Both *A* and *AB* are thus possible signal meanings in the depicted situation. Since the agent's I-language only allows for *A* to be expressed, the agent produces signal *X* to communicate speaker meaning *AB*. Signal *X* is used innovatively because the speaker meaning it conveys in this situation (*AB*) deviates from its conventional meaning (*A*).

The situation shown in (8) exemplifies the other aspect of pragmatic plasticity: innovative use through overspecification.

(8) *Overspecification*

Input:	I-language:	$X \rightarrow AC$
	Speaker meaning:	*A*
	Inferable meaning:	–
	Ignorable meaning:	*C*
Calculations:	Possible signal meanings:	*A*, *AC*
	Expressible signal meaning:	*AC*
Output:	Signal:	*X*

In this example, the agent needs to convey a speaker meaning *A* but his I-language only contains a construction for the more specific meaning *AC*. The agent can use this construction nonetheless as he knows that the meaning component *C* will be ignored in the present context. Like underspecification, overspecification constitutes innovative use because signal meaning (*AC*) and speaker meaning (*A*) differ from each other.

More complex communicative acts can, of course, combine all of the above forms of use. Consider example (9).

(9) *Various forms of use combined*

Input:	I-language:	$X \rightarrow AC$
		$Y \rightarrow D$
	Speaker meaning:	*ABD*
	Inferable meaning:	*B*
	Ignorable meaning:	*C*
Calculations:	Possible signal meanings:	*AD, ACD, ABD, ABCD*
	Expressible signal meaning:	*ACD*
Output:	Signal:	*XY*

In the situation shown in (9), the agent needs to convey a speaker meaning *ABD*. Of the three semantic units contained in the speaker meaning, two necessarily need to be specified (*A* and *D*) whereas one (*B*) can be left underspecified because it can be inferred from the context. Any signal meaning must therefore at least contain *A* and *D*. Furthermore, the agent knows that meaning *C* will be ignored in the present context because it is irrelevant. It can thus be added to a potential signal meaning without being an impediment to utterance interpretation. The only potential signal meaning that the agent can express with his I-language is *ACD*. He therefore produces the signal *XY*. The produced signal *combines* two symbols: $X \rightarrow AC$ and $Y \rightarrow D$. One of them, *Y*, is used *conventionally*, the other one, *X*, exhibits *overspecification*: only one of its two semantic units actually contributes to the speaker meaning. Finally, the signal meaning as a whole *underspecifies* the speaker meaning because it does not express *B* but rather leaves it to be inferred from the context.

There may be iterations in a simulation in which no signal is available to convey the intended meaning. An example of such a case is given in (10). In the depicted situation, the agent's I-language cannot express any of the possible signal meanings. Communication is thus not possible.

(10) *No signal available*

Input:	I-language:	$X \rightarrow AB$
		$Y \rightarrow D$
	Speaker meaning:	A
	Inferable meaning:	–
	Ignorable meaning:	C
Calculations:	Possible signal meanings:	A, AC
	Expressible signal meaning:	–
Output:	Signal:	–

Other models employ a process of random invention in situations where the I-language of their agents is not capable of expressing the desired meaning (cf. Hurford 2002 for an overview). In contrast to the models discussed in Hurford (2002), the model introduced here does not presuppose that speakers always manage to get their message across. It seems more plausible to assume that sometimes, humans do not have an appropriate expression to convey the concept they are thinking of, and that in such situations they do not invent a random new form for that concept. Therefore, the simulation simply moves on to the next iteration and provides the agent with a new speaker meaning and a new context if such a situation occurs.[7]

4.1.2.3 *Cases of synonymy*

In all examples that I have discussed so far, only one signal was available that could communicate the intended signal meaning in the given context. In many contexts, however, the agent has more than one option. We can distinguish between cases of what I will call *grammar-related synonymy* and cases of *context-related synonymy*. Grammar-related synonymy occurs if the agent's I-language, his grammar, provides more than one way of expressing a certain signal meaning. Context-related synonymy denotes situations where, in the given context, more than one expressible signal meaning can convey the intended speaker meaning.

Example (11) shows a case of grammar-related synonymy.

[7] In section 5.2.2, I will show how the success rate of an I-language changes over time.

(11) *Grammar-related synonymy*

Input:	I-language:	$X \rightarrow A$
		$Y \rightarrow B$
		$Z \rightarrow AB$
	Speaker meaning:	*AB*
	Inferable meaning:	–
	Ignorable meaning:	–
Calculations:	Possible signal meanings:	*AB*
	Expressible signal meaning:	*AB*
Output:	Signal:	*XY* or *Z*?

The described situation shows grammar-related synonymy because, even though only one signal meaning (*AB*) is possible, this signal meaning can be expressed in two different ways: either by producing *XY* or by producing *Z*.

In contrast, examples (12) and (14) depict situations that exhibit context-related synonymy. The first example shows a situation where underspecification is possible but not necessary:

(12) *Context-related synonymy: optional underspecification*

Input:	I-language:	$X \rightarrow A$
		$Y \rightarrow B$
	Speaker meaning:	*AB*
	Inferable meaning:	*B*
	Ignorable meaning:	–
Calculations:	Possible signal meanings:	*A*, *AB*
	Expressible signal meaning:	*A*, *AB*
Output:	Signal:	*X* or *XY*?

In example (12), the agent *can* but does not need to underspecify the intended speaker meaning *AB*: he can either convey it by employing signal *X*, whose signal meaning *A* would underspecify the speaker meaning, or by producing signal *XY*, whose signal meaning *AB* would fully specify the speaker meaning. This is a situation that is not uncommon in every-day language use. Consider, for instance, sentence (13). We can easily think of situations where it would not be necessary to repeat the adjective *old* but where it could be done anyway.

(13) Fifty old men and [old] women live in this residence.

A similar situation arises if overspecification is possible but not necessary, as shown in the following example:

(14) *Context-related synonymy: optional overspecification*

Input:	I-language:	$X \rightarrow AC$
		$Y \rightarrow A$
	Speaker meaning:	A
	Inferable meaning:	–
	Ignorable meaning:	C
Calculations:	Possible signal meanings:	A, AC
	Expressible signal meaning:	A, AC
Output:	Signal:	X or Y?

In example (14), the agent has the option to overspecify the speaker meaning but again is not confined to doing so: he can either convey the intended speaker meaning *A* by using signal *X*, which expresses the speaker meaning "literally," or he can employ signal *Y*, whose signal meaning contains the irrelevant component *C*, which will be ignored in the present context. Such a case occurs, for instance, when we have to decide whether we shall call Sally a chameleon, as in example (3) above, knowing that the irrelevant semantic components of the concept 'chameleon' will be ignored, or whether we shall be precise and simply state that she frequently changes her appearance.

In either case, grammar-related synonymy as well as context-related synonymy, more than one signal is available to convey the intended speaker meaning, and the agent needs to choose one of them. In reality, this choice may be highly situation-specific and depend on the social context of the communicative event: a poet, for instance, may choose a highly metaphorical signal that grossly overspecifies the intended speaker meaning, a firefighter who needs to communicate quickly will go for the shortest signal, and a scientist may try to be as explicit as possible.[8] These are just extreme cases of a more general tendency: signals are chosen for a whole range of reasons. Overall, the users' choices thus appear

[8]It may be argued, however, that extra meaning is conveyed thereby.

random and do not follow a clear-cut, consistent strategy. A neutral way of modelling the way users behave when they face a case of synonymy is therefore to have the agent choose a signal at random.

In contrast, if we were to assume a bias, we would introduce a bottleneck that would affect the outcome of the simulation (cf. Hurford 2002). In the following paragraphs, I will show that the two main contemporary pragmatic theories, neo-Gricean pragmatics and Relevance Theory, imply that there are indeed tendencies in speakers' behaviour that are more basic than the socially influenced decisions mentioned above. I will show how these biases can be included in the model. (Later, in chapter 5, I will analyse how the introduction of such a bias—as opposed to choosing a signal at random—influences the course of the simulation and what this can tell us about the cultural evolution of language.)

The first possible bias I want to discuss is based on *neo-Gricean* pragmatics. Horn (2004) introduces two principles that describe human communicative behaviour: the so-called Q-principle and R-principle. He argues that these two principles capture Grice's (1975) conversational maxims as well as Levinson's (2000) tripartite system.

(15) a. *Horn's Q-principle*
Make your contribution sufficient;
Say as much as you can (given the R-principle).
b. *Horn's R-principle*
Make your contribution necessary;
Say no more than you must (given the Q-principle).

In the model, Horn's Q-principle corresponds to the reasoning by which an agent determines what constitutes a possible signal meaning in a given context: a signal meaning has to provide sufficient information for the hearer to infer the speaker meaning. Given the Q-principle, the R-principle would then suggest that, if more than one possible signal meaning is expressible, the agent should choose the one that underspecifies the speaker meaning most: only what is absolutely necessary should be said explicitly.

The other major contemporary theory of pragmatics, Sperber and Wilson's (1995) *Relevance Theory*, describes the cognitive processes underlying communication in terms of the following two principles:

(16) a. *The communicative principle of relevance*
Every act of ostensive communication communicates a presumption of its own relevance.

b. *The cognitive principle of relevance*
Human cognition tends to be geared to the maximisation of relevance (that is, to achieving the greatest cognitive effect for the smallest possible processing effort.)

While the communicative principle of relevance describes the basic fact that signals are assumed to transfer information that is relevant, the cognitive principle of relevance deals with the actual interpretation of signals. It implies that if more than one possible signal meaning is expressible, the agent should choose the one that involves the smallest effort. (We can ignore the aspect of cognitive effect here because all possible signal meanings achieve the same cognitive effect, namely the inference of the intended speaker meaning.) The Relevance Theory literature deals with the goal of minimising the involved effort in two ways.

The first is based on the assertion that inference is cheaper than coding.[9] Carston (2002:289) states that "it follows from the processing effort side of the definition [...] that speakers should not be, and are not expected to be, as explicit as possible." From this observation, she concludes that "[t]hey should encode only what they cannot rely on their addressees to infer easily." This strategy corresponds to Horns R-principle. In terms of our computational model, it means that the agent should express the most underspecified signal meaning.

Another approach commonly found in the Relevance Theory literature describes processing effort in terms of accessibility: what is easier to access takes less effort. Carston (2004:822), as we have already seen, explains the process of utterance interpretation as follows: "[c]heck interpretive hypotheses in order of their accessibility—that is, follow a path of least effort until an interpretation that satisfies the expectation of relevance is found; then stop." In the same way, production can be viewed as a process by which the first accessible signal meaning is

[9]Note that there is an apparent contradiction: in chapter 3, I have shown that coding is simply a special case of inference; but if coding is inference, how can the former take more effort than the latter? The answer is that it is not the cognitive process of coding but rather the process of articulation, the process of actually producing the signal in the physical world, that takes more effort. Levinson (1995:95f.), for instance, identifies articulation as a "relatively slow and inefficient process, which acts as a bottleneck in the entire communicative procedure." Humans can think faster and with less effort than they can articulate or gesture. I will come back to this point later, in section 5.3.

expressed. This account is particularly plausible from a psychological point of view because it does not imply that the speaker first constructs all possible signal meanings and then selects one from among them: he simply produces whatever signal "comes to his mind first." In terms of the model, this means that the agent should choose the signal where the constructions involved in the production are most entrenched.

To summarise the discussion of how language use is implemented in the model, the following points can be made. The model's setup allows for use to be innovative because it makes the distinction between speaker meaning and signal meaning explicit and gives the agent the possibility to employ signals whose signal meanings under- and/or overspecify the speaker meaning he needs to communicate. In cases of synonymy, that is, where more than one signal is possible, the agents behaviour can be simulated in three ways: either (i) the agent is unbiased and chooses a signal meaning at random (this solution does not make any assumptions and its result therefore provides us with a NULL-hypothesis), or (ii) he is biased towards maximising underspecification, or (iii) he is biased towards maximising entrenchment. How entrenchment is realised in the model will be explained now.

4.1.3 Learning

An extant convention is used in an innovative way when it comes to exhibit pragmatic plasticity in a specific context, that is, when it is used in an under- and/or overspecified way: its form communicates a novel meaning. However, innovative use can only turn into an actual innovation if it is followed by a second process: learning. Without learning, the innovative usage of a linguistic form cannot persist beyond the instance of use in which it was created. Learning describes the process by which a usage event results in a change in the agent's I-language. In this section, I discuss how learning is realised in the computational model.

4.1.3.1 Entrenchment

As discussed in section 4.1.1, the present computational model assumes learning to be a faithful process. Consequently, the learning mechanism it employs is a simple one. It simulates the process of entrenchment: each form-meaning association activated in a usage event is entrenched further through that usage event. Activated form-meaning associations are (i) the constructions involved in the production of the signal, and (ii) the association between the produced form and the communicated meaning. Concretely, entrenchment is implemented as follows. Each construction in an agent's I-language is stored together with some usage-related information, namely (i) the number of times the construction has been used, and (ii) the number of the usage event in which the construction was last used. This information is updated after each usage event by (i) increasing the first value by one and by (ii) setting the second value to the number of the respective usage event. The former captures *frequency*, the latter *recency* of use. If a form-meaning association is not stored in the agent's I-language yet—alternatively, we may think of both of its values as being 0—it is added to the I-language with its number of use set to 1. Example (17) illustrates how the process of entrenchment is realised in the computational model.

(17)	I-language:	$X \rightarrow A$	(1, 12)
		$Y \rightarrow B$	(3, 10)
	Usage event no. 14:		
	Speaker meaning:	*ABC*	
	Inferable meaning:	*C*	
	Ignorable meaning:	–	
	Signal:	*XY*	
	Updated I-language:	$X \rightarrow A$	(2, 14)
		$Y \rightarrow B$	(4, 14)
		$XY \rightarrow ABC$	(1, 14)

The example shows an excerpt of an agent's grammar after the 13th usage event the agent has participated in. One of the shown constructions, which maps form *X* onto meaning *A*, has been used only once so far, namely in usage event no. 12. The other construction, mapping *Y* onto *B*, has been used 3 times, the last time in usage event no. 10. This is the grammar which the agent uses to convey the intended message *ABC* in usage event no. 14, which follows: the first construction is used to encode *A* and the second to encode *B*; meaning component *C* can be left underspecified in the given context. The agent thus produces *XY* as a signal and thereby communicates utterance meaning *ABC*. After this usage event, both extant constructions have been used once more and the respective values are increased by one. The value that indicates in which usage event the constructions have been used last is set to 14, the number of the current iteration. Additionally, the agent's I-language is enhanced with a new construction, namely one that associates the produced signal *XY* with the communicated meaning *ABC*. This new construction has only been used once, in usage event no. 14.

The degree of entrenchment of a form-meaning association can be calculated as the number of times the said association has been used, divided by the time for which it has not been used anymore. The entrenchment of a construction is thus the deeper, the more frequently and the more recently that construction has been used. Formula (18) summarises how entrenchment is measured in the computational model.

(18) $$e_{c,t} = \frac{n_{c,t}}{i_t - l_{c,t}}$$

The entrenchment e of a construction c at a time t is the number of usages of that construction at the given time ($n_{c,t}$), divided by the number of consecutive iterations in which it has not been used anymore—the latter being the difference between the total number of iterations at the time i_t and the number of the iteration in which the construction was used for the last time $l_{c,t}$.

4.1.3.2 Decay and loss

The entrenchment of a construction not only increases when it is used, it also decreases when the construction is not used. "Entrenchment is reinforced through use [...] and decays through lack of use (as any rusty second language learner can attest)" (Croft 2000:73). The phenomenon of decay is the reason why entrenchment cannot simply be represented as frequency of use in the model. Only by applying formula (18) do we get the effect that the entrenchment of a construction not only increases through use but also decreases through lack of use.

The current computational model allows the experimenter to define an entrenchment threshold. A construction is kept in the agent's memory as long as its entrenchment is above this threshold and is lost ("forgotten") otherwise. After each iteration, constructions that are not sufficiently entrenched anymore are removed from the agent's I-language. The entrenchment threshold can be set to any value between 0 and 1. If the threshold is set to 1, no construction is remembered at all; if the threshold is set to 0, none is ever forgotten. Thus, with the entrenchment threshold set to 0.5, a construction that has been used only once will be lost after it has not been used for two consecutive iterations. Likewise, a construction that has been used five times in total will be lost after it has not been used in ten consecutive usage events. This example is illustrated in (19).

(19)

Entrenchment threshold:	0.5			
Last usage event:	40			
I-language before loss:	$X \to A$	(1, 38)	Entrenchment:	0.5
	$Y \to B$	(5, 30)		0.5
	$Z \to C$	(3, 36)		0.75
I-language after loss:	$Z \to C$	(3, 36)		

The process of loss can either be understood as an actual decay of the entrenchment of constructions in an individual's memory, or it can be seen as a simulation of the effects that a generation turn-over would have. In a two-agent population, a generation turn-over would remove one agent and introduce a new one instead. The I-language of this new agent naturally differs from that of the remaining "old" agent: at the moment of its "birth," it is empty. The new agent's I-language will thus only ever contain constructions that were used after the point of its introduction into the simulation, and once the remaining old agent is removed in another generation turn-over, any constructions that have not been used since will be lost (see e.g. the discussions in Hurford 2002; Kirby 2002a). In the long run, the "survival" of a construction therefore depends on its frequency of use. This is exactly the same effect that loss has too: constructions that are not used frequently enough disappear from the language over time. There is no explicit generation turn-over in the model introduced in this chapter, but one can say that the mechanism of loss simulates the effects it would have. The function of the entrenchment threshold is roughly equivalent to that of the size of the learning bottleneck assumed in other models, that is, the number of usage events between one generation turn-over and the next (cf. Kirby 2002a).

4.1.3.3 Entrenchment as exemplar-based learning

It seems appropriate to briefly consider where the learning method employed in the described usage-based model is positioned within the field of learning theory. Learning through usage memorisation has been suggested in the literature under various names. In psychology, approaches to categorisation that incorporate it are known as exemplar-based models (Smith and Medin 1981; Barsalou 1990). Such models have been introduced to artificial intelligence in the form of case-based reasoning (CBR) systems (Stanfill and Waltz 1986; Bareiss et al. 1987; Koton 1988; Rissland et al. 1989) and have come to be known in the field of machine learning as instance-based learning methods (Aha et al. 1991;

Mitchell 1997). Memory-based approaches to natural language processing (Daelmans 1999; De Pauw 2003) are another instantiation of exemplar-based models. Memory-based learning is built on both instance-based methods of machine learning as well as the so-called analogical model of language developed by Skousen (1989, 1992).

Exemplar-based learning, as opposed to other models of learning, has two main characteristics. First, the learning process is simple: all that it does is store experienced usage events in the memory. No generalisations are made and no rules are induced during the learning process. De Pauw (2003:150) states that the knowledge acquisition phase of such models is trivial in comparison to other machine learning methods such as, for example, neural networks or probabilistic models: "no further knowledge in the form of induction rules, decision trees, etc. needs to be induced from the data, nor does the probabilistic distribution of the data need to be computed during or after the knowledge acquisition phase." The storage of exemplars in the memory is the minimum of which learning can consist. Aha (1997) consequently refers to this most simple type of learning as "lazy learning."

The second characteristic of exemplar-based models follows from this first one. Because they only store experienced usage events, "[i]nstance-based learning algorithms do *not* maintain a set of abstractions derived from specific instances" (Aha et al. 1991:1, emphasis in the original). Categorisation happens ad hoc in the process of use and not during learning. New usage situations are classified by finding the most *similar* exemplar in the memory and *extrapolating* its solution (De Pauw 2003:150). Batali (2002:116, emphasis in the original) summarises these two main characteristics of exemplar-based models as follows:

> [L]earners simply store all of their analyzed observations as *exemplars*. No rules or principles are induced from them. Instead, exemplars may be used intact, to express the exact meaning, or to interpret the signal, recorded in the exemplar. Exemplars may also be modified to construct new analyses of the mapping between a signal and a meaning.

Batali, in the above quote, mentions two types of use: conventional ("intact") and innovative ("modified") use. In conventional use, the new event is identical to a stored exemplar; in innovative use, the new event is only similar to a stored exemplar. As I have described in section 4.1.2, both types of use are also implemented in the introduced computational model. Exemplars are modified when they exhibit pragmatic plasticity, that is, when they come to under- and/or overspecify the communicated meaning in a specific context.

Examples (20) and (21) illustrate the difference between exemplar-based models and other models of learning. The sequence represented in (20) shows a learning method which generalises from the experienced data $X \rightarrow AB$ and $X \rightarrow AC$ to a new category $X \rightarrow A$.[10] In a second step, the knowledge thus acquired is then put to use without any change. In contrast, example (21) illustrates the exemplar-based paradigm. An experienced item $X \rightarrow AB$ is stored in the memory without any change. However, it is then used in a later situation in an overspecified way to express a more general concept. This new exemplar is stored again without any modification. Generalisation is the product of overspecified use here, and not something the learning algorithm does during knowledge acquisition. The more general concept $X \rightarrow A$ is then later used in an underspecified way and gives thus rise to a category that is more specific again: $X \rightarrow AC$.

(20)	E-language:	$X \rightarrow AB$	
		$X \rightarrow AC$	
		$\downarrow$	Learning (generalisation)
	I-language:	$X \rightarrow A$	
		$\downarrow$	Use (exact specification)
	E-language:	$X \rightarrow A$	

(21)	E-language:	$X \rightarrow AB$	
		$\downarrow$	Learning (storage)
	I-language:	$X \rightarrow AB$	
		$\downarrow$	Use (overspecification)
	E-language:	$X \rightarrow A$	
		$\downarrow$	Learning (storage)
	I-language:	$X \rightarrow AB$	
		$X \rightarrow A$	
		$\downarrow$	Use (underspecification)
	E-language:	$X \rightarrow AC$	

[10] The experienced data $X \rightarrow AB$ and $X \rightarrow AC$ could also be stored during learning beside the newly generalised concept. They are not shown in the example I-language for the sake of simplicity.

The key point is this. In exemplar-based learning models, learning is faithful and use is innovative, whereas in other learning models, use is faithful—a mere one-to-one application of what has been learnt—and learning is innovative. This representation makes it evident that one can draw a direct line from learning theory to the study of cultural evolution. The different theoretical models of learning are reflected in the different views about the locus of innovation in models of cultural evolution which I have discussed in chapter 2 of this thesis and in section 4.1.1 above.

Given its much defended psychological plausibility (Medin and Schaffer 1978; Brooks 1978; Hintzman 1986; Nosofsky 1986; Skousen 1989; Chandler 1993), exemplar-based learning has been applied surprisingly sparsely in computational models that simulate the evolution of language. Most of them employ alternative learning methods, such as rule-based learning (e.g. Kirby 2001, 2002a; Zuidema 2003; Hoefler 2006b), Bayesian learning (Briscoe 2005; Kirby et al. 2007), MDL-learning (Teal and Taylor 2000; Brighton and Kirby 2001; Roberts et al. 2005), or neural networks and connectionist learning (Christiansen and Devlin 1997; Batali 1998; Livingstone 2002; Smith 2003c). Batali (2002) presents the rare example of a model that applies an exemplar-based learning algorithm. Using an exemplar-based computational model, Batali studies how recursive grammars can emerge when agents negotiate a communication system. His model is evaluated and extended in Eddy (2005). The use of exemplar-based models in evolutionary linguistics gains further theoretical support by some linguistic evidence presented in Wray (2005). Wray argues, mainly on the basis of evidence from idiomatic or formulaic language, that exemplar-based models are more accurate descriptions of linguistic knowledge acquisition than models based on the induction of rules. A similar conclusion has been reached, on the basis of psycholinguistic evidence, by Chandler (1993). I conclude that a strong case has to be made for the consideration of exemplar-based learning models in the study of language evolution—be it computational or theoretical—because of the high psychological plausibility of such models. The computational model introduced in this chapter makes a contribution to this direction of research.

4.2 Integrating non-symbolic communication

So far, I have described a computational model that simulates iterated *symbolic* communication. But, as I have pointed out before, a model of language evolution ideally starts from a non-linguistic (and therefore non-symbolic) state. The question I therefore want to address now is how *non-symbolic* communication—and the *transition* from non-symbolic to symbolic communication—can be added to the computational model. To this aim, one must first be clear about (i) what properties symbolic and non-symbolic communication share, and (ii) what distinguishes the two forms of communication.

In chapter 3, I have shown that both symbolic and non-symbolic communication are ostensive-inferential. In ostensive-inferential communication, a signal provides information (the *signal meaning*) which is used as a *clue* to infer the intended speaker meaning from the context. The hypothesis I have established is that the processes of ostensive-inferential communication are the same in both symbolic and non-symbolic communication. This would entail, as a derived hypothesis, that one should be able to represent non-symbolic communication in the introduced computational model *without* having to modify the mechanisms of use and learning that the model employs for symbolic communication. In the remainder of this chapter, I will show how (and why) this can be done.

But what, then, is the difference between symbolic and non-symbolic communication? In short, as opposed to symbolic communication, non-symbolic communication employs as signals forms that are not conventionally associated with a meaning (yet)—or at least does not allude to such conventions if they exist. The problem now is that the introduced mechanisms of use depend on the notion of a signal meaning being provided, namely in the form of a meaning that is conventionally associated with the produced signal. If we do not want to modify the employed mechanisms of use and learning, our proposal of how non-symbolic communication (and the transition to symbolic communication) can be represented must therefore answer the following question: what signal meaning is provided by forms that are *not* conventionally associated with a meaning—and how can they be represented in the introduced computational model? I will begin to answer this question by asking what forms are in the first place.

4.2.1 What are forms?

The thesis I want to put forward here is one that may at first sound paradoxical, namely that *forms are meanings*. To this aim, I will briefly review what meanings are and then continue to explain why forms are meanings.

The representation of I-language in the present model is compatible with the way it is described by construction-based approaches to grammar. There are various versions of construction grammar (see e.g. Croft and Cruse 2004:ch. 10 for an overview) but they all share the basic characteristics I have worked out in 2.2.3.2. In the following discussion, I will mainly use the terminology and concepts suggested by Langacker (1987:ch. 2) for his instantiation of a construction grammar: Cognitive Grammar. Langacker's account is particularly appealing for several reasons. First, it constitutes one of the foundational and most influential approaches to language within the framework of cognitive-functional linguistics. Second, Langacker addresses an extensive range of fundamental issues relevant to the modelling of grammar in a particularly accessible way—if one is ready to acquire his terminology and concepts, that is. Third, and probably most importantly, he makes a great effort to ground the concepts used in Cognitive Grammar in domain-general cognitive capacities; his descriptions are thus particularly apt to contribute to bridging the gap between no language and language, which must remain the ultimate goal of any study of language evolution.

4.2.1.1 Meanings

The first thing to be said about meanings is that they are *conceptual*. Meanings are conceptualisations of certain states of affairs in the world, for example, the concept of a unicorn, or the concept of a horn, or the concept of something being white, or the idea of somebody running, or the information that something is on top of something else, that a goal has been achieved by means of a certain behaviour, or the knowledge that something has a certain property. Langacker (1987:76) describes *semantic space*, that is, the space of all possible meanings, as "the multifaceted field of conceptual potential within which thought and conceptualization unfold" and concludes that "a semantic structure [i.e. a meaning] can then be characterized as a location or a configuration in semantic space."

The way in which meanings have been represented in existing computational models is determined by two factors. One is the aim of the model. Depending on what phenomena the model is designed to investigate, different levels of abstraction make sense in the representation of its meanings. If a level of abstraction is chosen that does not simply treat meanings as atomic units or black boxes but assigns them some internal structure, the researcher's theoretical assumptions about linguistic meaning might come to bear as a second influencing factor. Existing computational models have represented structured meanings in some variant of first-order logic (Hurford 2000; Kirby 2000, 2002a; Batali 2002), as vector matrices (Steels 1996a; Kirby 2001; Brighton and Kirby 2001; Brighton

2002; Smith 2003c) or as discrimination trees (Steels 1996b, 1997, 1999; Steels and Kaplan 2002; Smith 2003b, 2005a). Other models have not specified any semantic structure and treated their meanings as atomic (Smith 2004).[11] The general aim of any computational model must be to employ a representation that is kept as abstract as possible and is only as specific as necessary.

For the purposes of the current model, meanings could not be represented as atomic entities: we have seen that the processes of under- and overspecification, which are at the heart of the model, require the recognition of individual meaning components. On the other hand, it is not necessary to specify what these components are any further. The representation of meaning employed in the model reflects the fact that meanings can be composed of smaller semantic units but does not build in any further assumptions about the nature of these semantic units. As we have already seen, semantic units are simply represented as uppercase characters (*A, B, C, ...*), and meanings are modelled as sets of such units. The set of all semantic units that are available in a simulation thus defines the semantic space in which the agent's conceptualisations can unfold.

The psychological underpinnings for the conception of meanings employed in the computational model are again provided by Langacker (1987:ch. 2). Langacker points out that meanings are semantic units, and that a unit can itself contain further semantic units. He understands a *unit* in psychological terms as "a structure that a speaker has mastered quite thoroughly, to the extent that he can employ it in largely automatic fashion, without having to focus his attention specifically on its individual parts" (Langacker 1987:57). Semantic units can thus be conceived as *established concepts*. Langacker (1987:58) provides the following example of how an assembly of semantic units can itself attain the status of a unit:

> In learning what a unicorn is, for example, a person pays explicit attention to its horselike character, to the fact that it has just a single horn, and to the location of this horn on its head. Once the notion has the status of a unit, he evokes these specifications as a familiar gestalt, and need not attend to them individually.

It is crucial for our undertaking that even though meanings can be seen as semantic units, their components can still remain accessible. Langacker (1987:59)

[11] Models which do not incorporate meanings at all exist too (Teal and Taylor 2000; Roberts et al. 2005; Hoefler 2006b).

emphasises that "[i]t is important to observe that when a complex structure coalesces into a unit, its sub-parts do not thereby cease to exist or be identifiable as substructures [...]." Without this fact, that is, if meanings only appeared as unanalysable blocks, underspecification and overspecification would not be possible: as we have seen above, these two mechanisms build on the interlocutors' ability to recognise individual parts within meanings, such as the recognition of the inferable parts of an intended speaker meaning or of the ignorable parts of a signal meaning.

On a similar note, it must be mentioned that the fact that the model works with elementary semantic units (represented as upper-case characters) does not imply that the existence of universal semantic primitives, as advocated, for instance, by Wierzbicka (1972, 1989), is assumed: the model's elementary semantic units can simply be understood as representing conceptual units that the agent is capable of identifying, without claiming that these conceptual units could not be broken down themselves into even smaller components. Langacker (1987:87) points out that his conception of meanings as decomposable semantic units makes no claim as to whether the smallest units of linguistic significance are necessarily primitives. He rather asserts that units, even though they require no constructive effort, must still be seen as relatively complex entities, and that the decision to stop decomposing them at a certain level of abstraction is ultimately always more or less arbitrary:

> Semantic units are defined relative to knowledge structures, which can be extremely complicated, even for units that are minimal for most linguistic purposes (e.g. in one of its senses *balk* presupposes extensive knowledge of baseball, but it is a minimal unit from the standpoint of symbolic relationships). Analysis of knowledge structures can be carried out to whatever delicacy is required by linguistic considerations, but the fundamental units uncovered in this way need not be specifically linguistic, nor is it necessary to assume that they ultimately reduce to a specified list of primitives. (Langacker 1987:87)

In summary, three general properties of meanings have been worked out here: (i) meanings are *conceptualisations*, (ii) psychologically, they have the status of *units*, and (iii) they are potentially *decomposable* into smaller units. The question we have to answer now is whether these characteristics also apply to forms.

4.2.1.2 *Why forms are meanings*

Forms are *conceptualisations* of certain types of producible behaviours (signals). One is easily tempted to think of a form-meaning association as a mapping between something external or physical (a signal) onto something internal or psychological (a meaning)—or vice versa. But in reality, a symbol maps the conceptualisation of one external entity (a form) onto the conceptualisation of another external entity (a meaning)—as Saussure (1916) pointed out, both forms and meanings are psychological. Producible complex sound waves, for instance, are conceptualised as phonemes. Gestures, in turn, can be conceptualised as combinations of notions such as 'up', 'down', 'fist', 'flat', etc. (cf. e.g. Stokoe 1960, 1978). And, as I have already pointed out in section 3.3.1, a form can also be the concept that a behaviour X is carried out before a behaviour Y, which can be found in schematic constructions. In principle, there is no difference between (i) a meaning being associated with another meaning and (ii) a meaning being associated with a form. Take, for instance, the knowledge that movement often implies intention and the knowledge that the word *apple* conventionally denotes an apple. In both cases, the respective individual associates one concept with another concept; the difference only lies in the type of relationship that holds between the two associated concepts.

Langacker (1987:78) too makes the "seemingly contradictory but ultimately rather obvious proposition that sounds (at least for many linguistic purposes) are really concepts." He points out that "sounds are like other concepts, which normally involve coordinated specifications in various domains of semantic space" (Langacker 1987:78). In the same way in which the concept of an apple is derived from various perceptual inputs, sounds (or any other modification of the physical environment that serves as a signal) are conceptualised on the basis of respective perceptual inputs:

> Because language is a cognitive entity, the speech signal must be regarded not just in physical but in *psychological* terms; the cognitive representation of linguistic expressions derives most directly from auditory impressions, and only indirectly from the sound waves that give rise to these impressions. (Langacker 1987:78, emphasis in the original)

Langacker (1987:80) summarises the argument in the following way: "[m]eanings are conceptual entities, so the conceptualization of a sound can be considered a meaning." In other words: forms are meanings too.

Like other concepts, forms are psychological *units*: if they are sufficiently entrenched, they can be identified and produced without constructional effort. A person, for instance, whose native tongue distinguishes between voiced and voiceless fricatives will be able to identify the two respective classes of sounds effortlessly because for this person, each of them has the psychological status of a unit. In contrast, a person whose native tongue does not make this distinction will have to put constructive effort into recognising it because here, the two sounds do not constitute distinct psychological units. (Langacker 1987:57 provides an analogue example with the capacity of a French and an English speaker of producing an *[ü]* sound.) Many phenomena of sound change (especially phonological reduction) can be seen as a result of the respective form having come to be treated more as a unit and less as a composition of individual components.

Like other concepts too, forms are *decomposable* into smaller units. While I may treat voiced and voiceless fricatives as two distinct units that can be identified and produced in an automatised fashion, I will still be able to recognise that the two units have a common property (being a fricative) and a property that distinguishes them (being voiced or voiceless respectively). A gesture, even though it may be carried out effortlessly, can still be analysed, for instance, in terms of the type of movement it performed, the shape of the hand at the beginning, during and at the end of the gesture, or whether it was carried out fast or slowly.

In summary, we can state that forms are meanings in the sense that they are conceptual units that are potentially decomposable into smaller conceptual units. From this fact, one must conclude that any concept (meaning) for which direct evidence can be produced can theoretically become a form, and, conversely, that any concept that is used as a form can possibly also be the information that is "meant" in a communicative exchange. The *phonological space*, that is, the space of all possible forms, is thus not distinct from the semantic space but rather a subregion of it. Langacker (1987:78) consequently defines a linguistic symbol as "a correspondence between two structures in semantic space (broadly conceived), where one of the two occupies the phonological subregion in particular." In the computational model, forms are thus represented just like meanings as sets of upper-case characters that stand for elementary semantic units.

4.2.1.3 Iconicity

Models of language evolution usually do not acknowledge the fact that forms and meanings are "made of the same stuff:" different forms of representation are chosen for forms on the one side and meanings on the other side. One consequence of this design decision is that such models cannot represent the phenomenon of iconicity, that is, they cannot model situations in which the form resembles the meaning it is associated with. Iconicity is not only very frequent in language (see e.g. Haiman 1980b, 1983, 2008; Givón 1990; Simone 1995; Keller 1998; Nänny and Fischer 1999) but can also be assumed to have played a vital, albeit not exclusive, role in the emergence of many linguistic conventions (cf. section 3.2.2; Keller 1998). It therefore seems appropriate that a computational model that simulates language evolution can deal with this phenomenon. This is the case for the model introduced in this chapter—precisely because it properly reflects the fact that forms are conceptual entities just like meanings too.

If we want to know what it means for a form and a meaning to resemble each other, which is how iconicity is defined, we must ask how the similarity between two concepts can be described. Two concepts are similar if *some* but not *all* of their properties are the same. If all properties were the same, we would speak of identity; if none were shared, the two concepts would be distinct. A form-meaning association is therefore iconic if some semantic unit(s) contained in the meaning also occurs in the form. Example (22) shows such an iconic relationship between a form and a meaning as it may occur in the simulations: the semantic unit *X* is contained in the form on the left-hand side as well as in the meaning on the right-hand side.

(22) $XY \rightarrow AXB$

The representation employed in the present computational model can be used to capture any sort of iconicity at an abstract level. Imagine, for example, that someone paints a small yellow circle to signify the sun. The relation between the sign and the signified object is iconic: the two resemble each other. Example (23) shows how this scenario can be represented in the notation employed by the computational model. Some of the properties are shared by both the drawing and the sun (both are round and both are yellow-ish), others are specific to the icon's form (that it is a drawing or that it is small) or the icon's meaning (that the signified object is a celestial body or that it is hot).

(23) a. C = *celestial_body(x)*
D = *drawing(x)*
H = *hot(x)*
R = *round(x)*
S = *small(x)*
Y = *yellow(x)*

b. Icon: $DRSY \rightarrow CHRY$

Similar examples could be constructed for other instances of iconicity. The point to remember is that the possibility to represent such phenomena in a computational model is a consequence of recognising that forms are meanings too. This insight about the nature of forms can now be applied to our main problem, namely how non-symbolic communication can be modelled.

4.2.2 Forms as signal meanings

In returning to the question of how non-symbolic communication can be integrated in the introduced computational model, it is important to remember that both symbolic and non-symbolic communication are forms of ostensive-inferential communication and that a unified model would thus only apply one set of mechanisms of use and learning for both of them. The introduced model simulates use as the process by which a signal meaning develops into a speaker meaning—possibly by adding additional inferable information to the signal meaning and/or by removing ignorable aspects of it. In symbolic communication, the role of the signal meaning is played by a meaning that is conventionally associated with the produced form. Non-symbolic communication by definition does not make use of conventions. Thus, if we want to retain the developed mechanisms of use, we have to determine where the signal meaning in non-symbolic communication comes from.

4.2.2.1 Representing the information contained in cues

We can define signal meaning as the information that the production of a cue adds to the common ground of the interlocutors. In section 3.1.3, I have shown that it is the addition of this information that consequently triggers and enables the inference of the speaker meaning from the context. The computational model

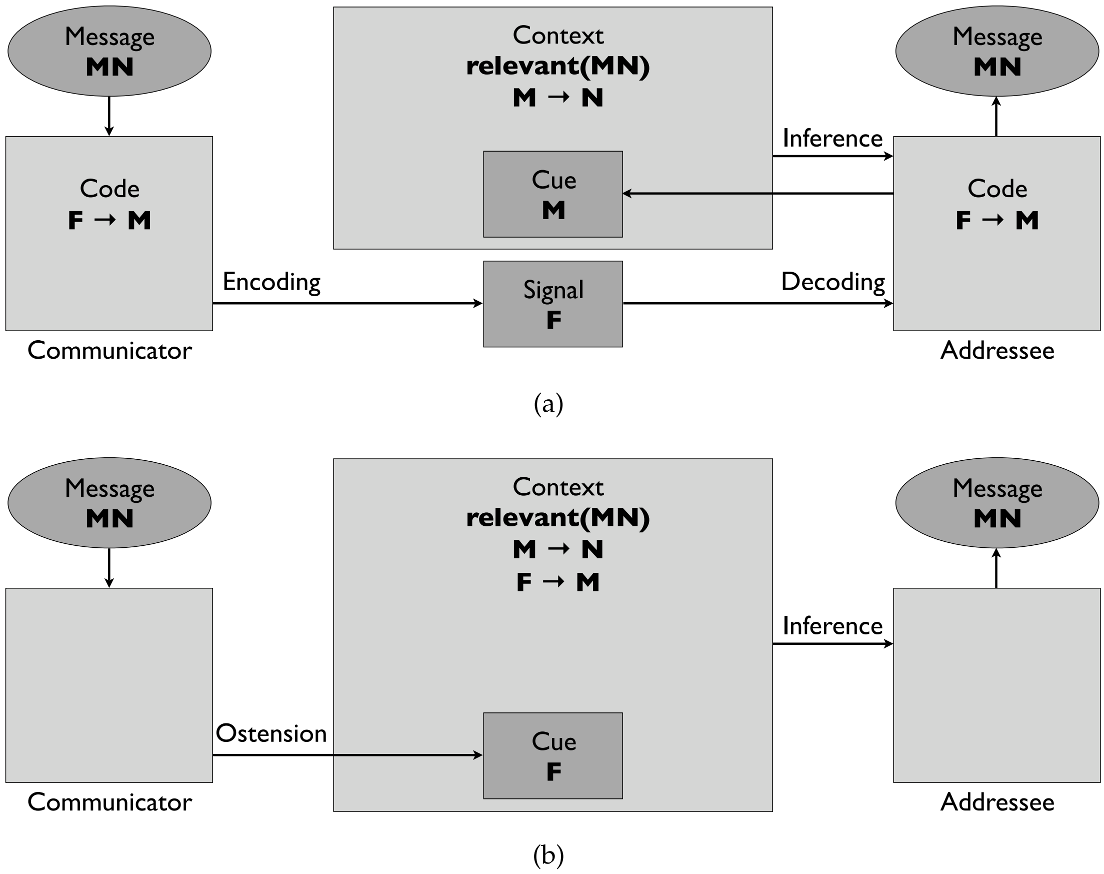


Figure 4.7: Modelling symbolic communication. The two figures depict the mechanisms of symbolic communication at different levels of abstraction. (a) Symbolic communication as consisting of two processes: coding and inference. This is the level of abstraction implemented in the computational model. (b) Symbolic communication as a special case of ostensive-inferential communication. This lower level of abstraction reflects the fact that coding is an inferential process too.

is formulated at the level of abstraction represented by Fig. 4.7(a), which I originally introduced in section 3.1.4. The information that a form F adds to the context is its conventional meaning M. At this level of abstraction, we can therefore interpret the implication $F \rightarrow M$ as a function that tells us that the production of F adds the information (i.e. signal meaning) M to the context. However, I also argued in section 3.1.4 that, at a lower level of abstraction, the processes described by Fig. 4.7(a) actually need to be understood as represented in Fig. 4.7(b). By producing a form F as a cue, the communicator adds information to the common ground she shares with the addressee. The addressee uses this information to draw inferences from context until he arrives at a relevant interpretation of the communicator's informative intention. (His recognition of her communicative intention is presupposed here.) To arrive at an interpretation, the addressee

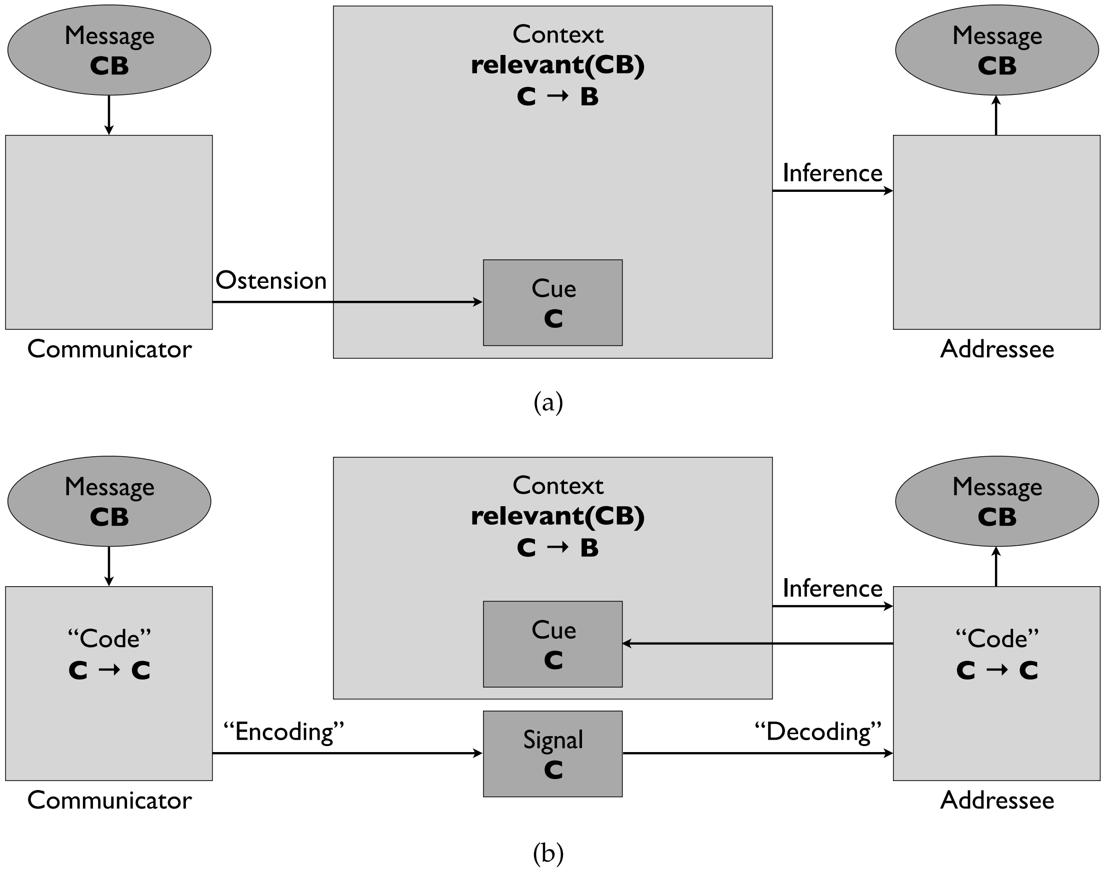


Figure 4.8: Modelling non-symbolic communication. The two figures depict the mechanisms of non-symbolic communication at different levels of abstraction. (a) Non-symbolic communication as described by the ostensive-inferential model of communication. The example shows the production of a coin (*C*) as a cue to communicate that one has the money for the bus (*B*). (b) This second figure shows how non-symbolic communication as depicted in (a) can be represented by the introduced computational model without any change to the mechanisms of use that this model employs, namely be equipping the agent with the knowledge that the production of any producible cue (in the example, the produced cue is *C*) will add the information that that cue has been produced to the common ground.

makes use of any sort of contextual knowledge, both symbolic ($F \rightarrow M$) and non-symbolic ($M \rightarrow N$). This process remains the same if no symbolic conventions are made use of in the interpretation of the provided cue. Fig. 4.8(a) represents such a case of non-symbolic communication. It shows the example, described in section 3.1.3, where a communicator produces a coin (*C*) as a cue to communicate that he has the money he needs for the bus (*B*). No conventional form-meaning association is made use of in this case of ostensive-inferential communication. The question now is how we can represent this example at the level of abstraction assumed in the computational model.

The solution can be found in our interpretation of the implication $F \rightarrow M$ as a function that tells us that the production of F adds the information (i.e. signal meaning) M to the context. In Fig. 4.8(a), the information that the cue C adds to the context is C itself. We can consequently represent this fact as $C \rightarrow C$: the production of C adds the information (i.e. signal meaning) C to the context. At first, this notation may appear tautological. However, we must be aware of the fact that the left-hand side of the implication stands for the produced cue whereas the right-hand side denotes the information that the production of the cue adds to the interlocutors' common ground. Any form F, upon production (and perception), always adds at least information F to the common ground—after all, this is the reason for producing it in the first place (cf. section 3.1.3). The immediate information that an addressee recovers when a communicator produces some form F as a communicative stimulus is the information that the respective form has been produced. On the basis of this first bit of information, he might infer other information—for example, if he knows that the produced form is associated with a conventional meaning. However, even in the absence of a convention, he will still employ the information that F was produced as the basis for the inference of the speaker meaning from context. This fact is what the implication $F \rightarrow F$ expresses.

By adopting this notation, we gain the means to represent the example of non-symbolic communication depicted in Fig.4.8(a) at the level of abstraction assumed in the computational model without having to change the involved mechanisms of use and learning, namely as shown in Fig. 4.8(b). The minimal "knowledge" that a communicator will always have is the knowledge that any producible behaviour will add the information that the respective behaviour was produced to the common ground he shares with a potential addressee. In the depicted example, the agent knows that (i) he can produce C, and that (ii) if he does produce C, this information will be added to the context.[12]

4.2.2.2 Self-symbolisation

The introduced notion of $F \rightarrow F$ can be seen as a limiting case of iconicity, where not only some but all properties of a form and a meaning associated with it are the same. Langacker (1987) states that in cases which I have analysed as a form

[12]Note that I only label this knowledge as "code" (NB in scare quotes) to indicate that it is stored in the same way as symbolic conventions in the computational model.

serving as signal meaning, the form is used as a symbol for itself.[13] He explains that from a theoretical point of view, such cases are nothing else than "an expected limiting case in the spectrum of conceivable symbolic relationships" and justifies his analysis by pointing out that "nothing in the definition of a linguistic symbol—a correspondence between two structures in semantic space, one of them phonological—inherently rules out the possibility that the two structures might happen to be identical, i.e. that a phonological structure might symbolize itself" (Langacker 1987:79ff.). To illustrate his claim, Langacker (1987:80) provides an example that contains both symbolic and non-symbolic aspects:

(24) The boy went [NOISE].

In sentence (24), the signal meaning is made up of the conventional meaning of the form *the boy went X* as well as the conceptual properties of the form *[NOISE]*. Instead of providing a symbolic description of the noise that the boy produced, the speaker makes a noise that has similar (or possibly identical) conceptual properties. The linguistic part of the sentence thus represents a symbolic clue (the form's conventional meaning functions as signal meaning), whereas the non-linguistic part constitutes a non-symbolic clue (the form itself serves as signal meaning).

4.2.2.3 Simulating non-symbolic communication

In summary, the idea of self-symbolisation provides us with the notational possibility to represent a form alone just like a conventional form-meaning association. It allows us to simulate non-symbolic communication in the introduced computational model without having to modify the employed processes of use and learning. The mechanisms of use implemented in the computational model will treat the right-hand sides of these mappings as the signal meanings elicited by the forms on the left-hand side, irrespective of whether form and meaning are identical or not. Example (25) illustrates how non-symbolic communication is simulated in the model. It shows a situation where the agent does not have any conventional form-meaning associations in his I-language: he only possesses the capacity to produce a signal *AC* which is not associated with any conventional meaning (yet).

[13] Note that Langacker uses the term "symbolise" in a more general sense than I have in this chapter. My definition includes the notion of conventionality whereas his seems to be roughly equivalent to the more general concepts of signifying or denoting. In this sense, the term "self-symbolisation" may be somewhat misleading and should rather be "self-signification."

(25) *Non-symbolic communication*

Input:	I-language:	$AC \rightarrow AC$
	Speaker meaning:	AB
	Inferable meaning:	B
	Ignorable meaning:	C
Calculations:	Possible signal meanings:	A, AB, AC, ABC
	Expressible signal meaning:	AC
Output:	Signal:	AC
	Updated I-language:	$AC \rightarrow AC$
		$AC \rightarrow AB$ (new)

In the example, the agent needs to express a speaker meaning *AB*. He also knows that the semantic unit *B* can be inferred from the context once *A* is given, and that the semantic unit *C* would be ignored if it was part of the signal meaning. The agent can convey the intended speaker meaning by producing the signal *AC*, which provides direct evidence (cf. section 3.1.3.2) for one of the possible signal meanings. This signal meaning at the same time under- and overspecifies the speaker meaning: *B* needs to be inferred from context and *C* can be ignored.

After this usage event, a new association between the employed signal (*AC*) and the communicated meaning (*AB*) is added to the agent's I-language. This is the first conventional form-meaning mapping that the agent comes to possess, and it can later be re-used in an instance of symbolic communication. The under- and overspecified use of a non-symbolic signal thus initiates the transition from non-symbolic to symbolic communication. Note that this addition of the conventional form-meaning mapping also introduces a first situation of polysemy: the form *AC* can now be used to signify either itself (*AC*) or a conventional meaning (*AB*) that is different from it. One can thus make the general statement that every form, every concept for which direct evidence can be produced, has the capacity to signify at least (i) itself, and potentially (ii) one or more conventional meanings.

4.2.2.4 *The initial state of a simulation*

In the initial state of a simulation, the agent does not possess any conventional form-meaning association in his I-language. All he is equipped with is a set of producible forms. The sum of these forms constitutes the phonological space available to the agent, which is itself a subset of the semantic space, the sum of all concepts that the agent could possibly "mean" in communicative situations. Such an initial state of a simulation is exemplified in (26). Note that the semantic space and the phonological space within which a simulation can unfold are defined by the experimenter: the former explicitly by providing a set of elementary semantic units (from which speaker meanings and context are randomly selected in every iteration; see appendix A), the latter implicitly by equipping the agent with an initial set of producible forms.

(26) *Initial state of a simulation*

Semantic space:	all subsets of $\{A, B, C, ..., X, Y, Z\}$
Initial I-language	
(producible forms):	$X \rightarrow X$
	$Y \rightarrow Y$
	$AZ \rightarrow AZ$
Phonological space:	all subsets of $\{X, Y, AZ\}$

What are the conditions under which an agent will be able to use such an initial set of producible forms to convey some speaker meaning? One possibility is that some of the available forms provide *direct evidence* for the speaker meaning the agent needs to convey—a process which I have discussed in section 3.1.3.2. In such a case, the agent can express the full speaker meaning explicitly by producing some of the available forms, as illustrated in example (27).[14] Speaker meaning (X), signal meaning (X) and the form of the signal (X) are identical. However, non-symbolic communication by provision of direct evidence does not yield a new, conventional form-meaning association (i.e. one where form and meaning are not identical); it only entrenches existing forms as units.

[14]Note again that the recognition of the produced behaviour as a communicative stimulus is presupposed by the model.

(27) *Production of direct evidence*

Input:	I-language:	$X \rightarrow X$
		$Y \rightarrow Y$
		$AZ \rightarrow AZ$
	Speaker meaning:	X
	Inferable meaning:	–
	Ignorable meaning:	–
Calculations:	Possible signal meanings:	X
	Expressible signal meaning:	X
Output:	Signal:	X
	Updated I-language:	$X \rightarrow X$
		$Y \rightarrow Y$
		$AZ \rightarrow AZ$

A transition from non-symbolic to symbolic communication can only occur in the model if some initially available form exhibits pragmatic plasticity in use. Like this, the model simulates both iconic as well as non-iconic cases of non-symbolic communication. An iconic form-meaning association arises if the produced form is not identical with the conveyed speaker meaning (as it is the case when direct evidence is provided) but only resembles it. Such a situation occurs if some (but not all) elements of the speaker meaning are inferred and/or if some (but not all) elements of the signal are ignored, that is, in cases where the produced form partially under- and/or overspecifies the speaker meaning. An example of such a case is shown in (28). The new, conventional form-meaning association that emerges from the depicted instance of non-symbolic communication ($AZ \rightarrow AB$) is iconic: both form and meaning contain the semantic unit A.

(28) *Partial under- and overspecification*

Input:	I-language:	$X \rightarrow X$
		$Y \rightarrow Y$
		$AZ \rightarrow AZ$
	Speaker meaning:	AB
	Inferable meaning:	B
	Ignorable meaning:	Z
Calculations:	Possible signal meanings:	A, AB, AZ, ABZ
	Expressible signal meaning:	AZ
Output:	Signal:	AZ
	Updated I-language:	$X \rightarrow X$
		$Y \rightarrow Y$
		$AZ \rightarrow AZ$
		$AZ \rightarrow AB$ (new)

However, the model also simulates cases where the relationship between an employed signal and the communicated meaning is not iconic. The model simulates such cases of non-symbolic communication as shown in example (29). In the depicted situation, the *whole* speaker meaning (B) can be inferred once the speaker has made manifest his communicative intention by producing a stimulus (Y) that is recognised as such in the present context. The *whole* signal meaning (Y) of the signal itself is ignored. The signal meaning Y thus completely under- and overspecifies the speaker meaning B. The result is the addition of a conventional form-meaning association where form and meaning do not resemble each other, that is, where the relationship between the two is not iconic.

(29) *Complete under- and overspecification*

Input:	I-language:	$X \rightarrow X$
		$Y \rightarrow Y$
		$AZ \rightarrow AZ$
	Speaker meaning:	B
	Inferable meaning:	B
	Ignorable meaning:	Y
Calculations:	Possible signal meanings:	B, Y, BY
	Expressible signal meaning:	Y
Output:	Signal:	Y
	Updated I-language:	$X \rightarrow X$
		$Y \rightarrow Y$
		$AZ \rightarrow AZ$
		$Y \rightarrow B$ (new)

In summary, one can state the following points. While the production of direct evidence does only further entrench an existing signal, any use of under- and/or overspecification, any case in which the employed signal exhibits pragmatic plasticity, results in the addition of a conventional form-meaning association, a mapping where a form is associated with a meaning that is different from its own conceptual properties. Symbolic communication, communication that employs conventional meanings as clues, will thus emerge in the simulation if situations occur where the form of an available non-symbolic signal is used to under- and/or overspecify an intended speaker meaning. Note that the employed processes of under- and/or overspecification do not distinguish between a non-symbolic signal (i.e. a form that is used to convey its own conceptual properties) and a symbol (i.e. a conventional form-meaning mapping). Non-symbolic signals will continue to be employed if they help convey a speaker meaning even when symbols are already available in other cases. The computational model introduced in this chapter can therefore not only simulate the transition from non-symbolic to symbolic communication but also the co-existence and combined use of both these two forms of ostensive-inferential communication.

4.3 Conclusion

In this chapter, I have introduced a computational model that is based on the theoretical considerations presented in the previous chapters of the thesis. While most current computational models of language evolution study the effect of imperfect learning and represent use as a faithful process, this model is designed to investigate the effect of innovative use and consequently makes the idealising assumption that learning is faithful. In general, the model simulates an individual's iterated involvement in communicative interactions and is aimed at studying the evolution of that individual's I-language in the course of this process.

One of the main questions addressed is how innovative use can be simulated. To this aim, I have made the following points. In order to represent innovative use, the model needs to reflect the distinction between signal meaning and speaker meaning. This has not been done in existing computational models of language evolution. I have argued that use is conventional if signal meaning and speaker meaning are the same; it is innovative if they differ. In the latter case, the signal exhibits pragmatic plasticity: the signal meaning under- and/or overspecifies the speaker meaning. In each iteration of a simulation, the agent is therefore presented with a randomly generated speaker meaning which he has to convey in an also randomly generated context. This context determines if and how the signal he employs can under- and/or overspecify the intended speaker meaning. In summary, I have modelled innovative use as pragmatic plasticity, as context-specific under- and/or overspecification of the speaker meaning. These mechanisms are kept at an abstract level, so that the introduced computational model can capture a wide range of innovative linguistic phenomena such as the ones listed in the introductory chapter of this thesis.

The model employs a maximally simple form of learning, exemplar-based learning, to simulate the mechanism of entrenchment. Although exemplar-based learning has a high psychological plausibility, it has only been applied very sparsely in existing computational models. This form of learning only stores observed data—in the case of the model the signal used and the communicated meaning—without making any generalisations. The model simulates the entrenchment of form-meaning associations so that it is enforced through use and decays through lack of use. Like this, infrequently used associations eventually get lost, which creates an effect similar to the one caused by the learning bottlenecks of models that, in contrast to this one, assume use to be faithful and learning to be innovative.

Unlike many other models, the computational model introduced here can simulate symbolic as well as non-symbolic communication. I have argued that forms and meanings are really made of the same stuff: both are conceptual entities. This fact is reflected in the way forms and meanings are represented in the model, namely as sets of elementary semantic units. One advantage of this representation is that iconicity can be modelled. Another one is that it allows the model to reflect the fact that forms themselves can serve as meanings and thus to simulate non-symbolic communication in the same fashion in which it simulates symbolic communication. I have argued that in symbolic communication, the conventional meaning of the signal serves as a clue (signal meaning) that triggers the inference of the speaker meaning from the context, whereas in non-symbolic communication, it is the form of the signal itself that performs this function. The same mechanisms of use can be employed in both cases because these mechanisms apply to any form-meaning association, be it the association between a form and its conventional meaning or the association between a form and its own conceptual properties (self-symbolisation). The computational model thus not only is capable of simulating the emergence of symbolic communication from an initial state where only non-symbolic communication is available but also accurately reflects the fact that symbolic and non-symbolic communication are both instances of ostensive-inferential communication and employ the same set of cognitive mechanisms.

In the next chapter, the computational model introduced here will be used to address certain aspects of the design puzzle.

CHAPTER 5

The design puzzle

The observation that language has the appearance of design for communication has proven defining for the course that the study of language evolution has taken. Besides the emergence puzzle, the design puzzle has become the main explanandum of evolutionary linguistics: how come language looks as if it had been designed for the purpose of communication? As I have already mentioned in the introduction to this thesis, the answers given to this question vary depending on whether language (or at least some core component of it) is assumed to be a specific innate module of the mind or not, and I will only briefly recapitulate the main points here. The former position yields a biological explanation for the design puzzle, the latter does not. Advocates of the innateness assumption, if they do acknowledge that language has the appearance of design (note that some very influential figures in the field, e.g. Chomsky, do not entertain this view) typically invoke natural selection as the only biological process that can bring about complex adaptive structures and therefore the appearance of design. This is the position advocated by Pinker and Bloom (1990) in their ground-breaking paper on *Natural language and natural selection*. If, as is the aim of this thesis, we do not want to make the innateness assumption, namely that the core of language is a genetically determined, domain-specific biological module of the mind, but rather presuppose that language is exclusively based on domain-general cognitive capacities, we cannot resort to biological explanations. As detailed in chapters 1 and 2, the processes of cultural evolution have been suggested as an alternative.

However, while models of the cultural evolution of language have shown that these processes can indeed yield complex design, they have typically focused on design for acquisition rather than design for communication (cf. sections 1.1.2, 2.3.2.1 and 4.1.1). The objective of this thesis is consequently to devise a model of linguistic cultural evolution in which the specifics of communication are made more explicit. To capture these specifics, I have proposed the notion of pragmatic plasticity. In chapter 3, I investigated the role of pragmatic plasticity with respect to the emergence puzzle. In this chapter, I will do the same for the design puzzle, and argue that pragmatic plasticity is one of the reasons why and how language has come to exhibit the appearance of design for communication. I will first address the question of what it actually means for something to exhibit design for communication (section 5.1). I will then look at two specific "design requirements" for a communication system such as language: expressivity (section 5.2) and signal economy (section 5.3). Finally, I will investigate what role pragmatic plasticity plays with regard to a putatively dysfunctional property of language: ambiguity (section 5.4). For each of these issues related to design for communication, I will provide a theoretical discussion as well as a set of simulations on the basis of the computational model introduced in the last chapter.

5.1 Design for communication

Before I can discuss the role of pragmatic plasticity with regard to the emergence of the appearance of design for communication, it is appropriate to define the concepts in question. The aim of this section is hence to work out what it actually means for something to have the appearance of design for communication. What conditions does a symbolic communication system need to fulfil, what characteristics does it have to exhibit, in order to have the appearance of design for communication? I will tackle this question in two steps. First, I will discuss the notion of design in general. I will ask what design is and how it can come about. In a second step, I will then apply the gained insights to communication and work out, at a general level, what design requirements a communication system needs to meet.

5.1.1 What is design?

The notion of design refers to the relationship between the *form* or properties of a cultural artefact or a biological trait and its *function*.[1] It describes the fact that, because of its properties, an artefact or trait is able to perform a certain

[1] The use of the term "form" in the sense of the totality of the properties of a cultural artefact or biological trait is not to be confused with its use to denote one half of a symbol.

function.[2] Design can thus only be understood in relation to a goal that needs to be achieved in a particular *environment*, and in relation to specific problems that the environment poses to achieving the goal. An artefact or trait exhibits design if its properties are *functional*, that is, if they provide a way of overcoming these problems. We can thus only judge if an entity exhibits design, or the appearance of design, if we are clear about what this design is *for*, that is, what *goal* it is meant to achieve and what *environment-inflicted obstacles* it overcomes to do so.

The two-fold definition of design or functionality that I have suggested here corresponds to a similar account provided by Hurford (1990:96). Hurford points out that for a tool—he mentions a spade and human language as examples—to exhibit design, it must accommodate a *purpose* as well as a *user*. The former corresponds to what I have referred to as goal or function; the latter is included in what I have described as the environment in which the goal needs to be achieved and which may pose specific problems to doing so. In this context, I agree with Kirby (1999:12f.), who, with regard to Hurford's distinction, emphasises that design can only be understood properly if both aspects are considered in a unified manner. Neither the goal alone nor the environment alone can define if an entity exhibits design.

Design, or the appearance of design, can be brought about by three processes: (i) biological evolution through natural selection, (ii) deliberate engineering, (iii) cultural evolution. In biology, the appearance of design is used as an indicator that a specific trait is an adaptation, that is, that it has been brought about by natural selection. Natural selection is viewed as the only biological process that can elicit complex design for some function. Therefore, if we acknowledge that language exhibits complex design and if we assume that it is a specific, genetically determined capacity, we must conclude that it is the product of (biological) evolution by natural selection (e.g. Pinker and Bloom 1990). It seems to me, however, that over the fascination that natural selection can bring about the appearance of design, the obvious is sometimes almost forgotten. *Actual* design, in its non-metaphorical sense, is the product of deliberate engineering. It is the result of a conscious, purpose-driven problem-solving process. Recent developments in genetic engineering aside, deliberate engineering does, of course, not produce biological entities but cultural artefacts. However, and this is the second obvious fact that is often overseen, most artefacts of our culture are the result not

[2] Kirby (1999:1), citing Cziko (1995), characterises the appearance of design as a "fit" of form and function and consequently interprets the design puzzle as a "puzzle of fit."

of a single, overarching problem-solving act of one individual but an accumulation of a myriad of much smaller problem-solving acts performed over a long period of time, that is, they are the product of cumulative cultural evolution as described in chapter 2 of this thesis. The solving of many small problems may eventually lead to an artefact that looks as if it had been designed to solve a much bigger problem. Only through such cumulative cultural evolution could artefacts of the complexity of aeroplanes, democratic institutions, computers, or language, which each, as a whole, exceed the innovative capacities of a single individual, have come about. However, it is not necessary—and probably not even possible—that the agents involved in the individual steps leading up to the cultural evolution of such an artefact are aware of the overarching problem the artefact is eventually going to solve. They may simply focus on providing ad-hoc fixes in the pursuit of smaller, momentary goals. After this brief characterisation of the notion of design and the processes by which it can be brought about, we can now turn to asking what it means for an entity to exhibit the appearance of design *for communication*.

5.1.2 What is design for communication?

In order to determine what constitutes design for communication, we first have to be clear about what communication is, in particular, what goal it has and what obstacles have to be overcome to achieve this goal. I have already discussed the foundations of communication in section 3.1, but it makes sense to recapitulate the main points of this discussion here from the point of view of design. The goal of communication is to transfer information from one individual's knowledge to another individual's knowledge—ultimately to manipulate that other individual's behaviour. An ideal communication system would thus allow humans

1. to communicate successfully
2. whenever they want (i.e. independently of the specific context)
3. whatever they want (i.e. all information that needs to be transferred)
4. at minimal cost (energy, time).

Note that I address the question of design for communication on a more general level than Pinker and Bloom (1990). They base their analysis on very specific assumptions about the nature of the transmitted information. I, on the other side, ask what design for communication means in general, independent of the assumed structure of the communicated information. From this perspective, design for communication seems to come down to the four requirements listed above.

The first problem that human communication needs to overcome is the fact that humans are not telepathic and that information cannot be transmitted from one mind to another *directly*. As I have already explained in section 3.1, the solution to this obstacle is for the communicator to modify the physical environment of the addressee in such a way that the addressee can infer the information the communicator wants to transmit. Ostensive-inferential communication, that is, the use of ostensive stimuli or cues, thus meets the first design requirement listed above. It ensures successful communication by providing a method of overcoming the problem, caused by the environment in which human communication is meant to happen, that one cannot directly manipulate one's conspecifics' minds.

However, the use of ostensive stimuli brings with it some restrictions: they do not allow for humans to transmit information "whenever they want." Rather, the use of ostensive cues is only possible in specific contexts, as the interpretation of such communicative stimuli crucially depends on the availability of inference from context. A better-designed communication system would therefore be one that allows for the transfer of information *independently of (situation-specific) context*. As I have already shown in section 3.2.1, the use of conventionalised form-meaning associations provides an answer to this design requirement. The emergence of conventions thus also brings about a further level in the appearance of design for communication.[3] At this point, we can further specify the original question and now ask by what properties a well-designed *conventionalised communication system* can be characterised.

A communication system that uses conventionalised form-meaning associations meets the first two requirements from the list above—it thus already exhibits the appearance of basic design for communication. However, such a system fulfils its communicative function better the closer it gets to meeting the second pair of requirements too. To achieve the first of these two remaining requirements, the *expressivity* of the conventionalised communication system must be maximised.

[3]A further aspect of the second design requirement, that communication be context-independent, which I will not further discuss here, is the medium chosen for the ostensive stimuli. In general, sound may be better designed to meet this requirement than gesture. Bickerton (2003:81) speculates that "[t]he original mixture of isolated grunts and gestures may have eventually settled on the vocal mode merely through the exigencies of communicating at night, over distance, or in dense vegetation." For an alternative approach, see Goldin-Meadow (2008).

The second remaining requirement, namely that cost must be minimised—Croft (2000:235) speaks of "the principle that speakers minimize the effort involved in their linguistic expression (Keller 1994)"—has given rise to several notions of economy or simplicity. We can distinguish between at least code simplicity, processing economy and signal economy. *Code simplicity*, which is sometimes also referred to as "grammar simplicity" or "representational economy"—has been given some attention in recent computational studies of language evolution (e.g. Brighton and Kirby 2001; Brighton 2003, 2005; Zuidema 2003; Roberts et al. 2005; Hoefler 2006b). Langacker (1977:23) specifies that "code simplicity pertains to the number of different fixed expressions, patterns, and locutions that a speaker must learn, remember, and manipulate in use." In this sense, a grammar that exhibits code simplicity appears to be particularly well designed for acquisition. Despite the fact that there may be certain empirical problems with the concept (I have pointed them out in section 2.3.3.1), code simplicity is central to many acquisition-driven models of language change and evolution (cf. section 2.3.2.1). However, since code simplicity seems to constitute a form of economy that is more closely linked to design for acquisition (and thus to notions such as the poverty of the stimulus and the learning bottleneck, which I do not deal with in this thesis) than to design for communication, I will not pursue it any further in this chapter. *Processing economy* has usually been interpreted in terms of parsing efficiency (e.g. Hawkins 1994, 2004; Kirby 1999) and has mainly been applied to syntactic phenomena and therefore been grounded, at least to some extent, in specific (generative) syntactic theories. In the introduced computational model, a grammar exhibits higher processing economy the deeper the average entrenchment of its conventions are (relative to the frequency with which they are used). This aspect of design for communication is maximised in the model if the agent consistently selects (and thus also further entrenches) the most entrenched constructions. However, the aspect of cost minimisation that I will look at more closely in this chapter is that of *signal economy*, that is, "economy in regard to production of the physical speech signal" (Langacker 1977:102f.). This aspect of simplicity, even though it is frequently invoked as a motivation for language change, has as yet been given surprisingly little consideration in studies of language evolution.

The remainder of this chapter will be dedicated to a discussion of how these two aspects of appearance of design for communication, expressivity and signal economy, can come about in the course of linguistic cultural evolution. I will show that pragmatic plasticity plays a central role in this process just as it is responsible for the emergence of symbolic communication, that is, communication by means of conventionalised form-meaning associations, in the first place.

Finally, the claim that language exhibits the appearance of design for communication is often countered by listing putatively dysfunctional[4] features (e.g. Chomsky 2002:106ff.). Pinker and Bloom (1990) deal with some of them[5], but they do not address the puzzle posed by ambiguity. My account of how expressivity and signal economy come about, in contrast, will be complemented by a discussion of the role that ambiguity plays in it. I will address this question in the last part of the chapter.

5.2 Expressivity

One of the most defining features of human language, as opposed to the communication systems of other animals, is its vast expressive power. Pinker (2003:16) calls it "[t]he most remarkable aspect of language." Similarly, Hurford (2004:554) speaks of a "huge quantitative gap" from non-human to human communication systems. He mentions the sheer difference in size between "the inventory of arbitrary symbols used by animals (up to about 30 distinct calls used by wild chimpanzees) and the vocabulary of human languages, which contain many tens of thousands of items." Hurford goes on to point out that there is not just a quantitative but also a qualitative difference: human vocabularies are learnt whereas animal calls are innate. Curiously, a lot of consideration has been given in evolutionary linguistics to the question of how humans can acquire such huge symbolic systems, and why other animals cannot, but little has been said about how such systems come about in the first place. The latter issue, it seems to me, deserves to be much more at the forefront of evolutionary linguists' interests. We must assume that a symbolic communication system started out small—ultimately with a (hypothetical) "first symbol." The question which arises then is how we can get from a small initial symbolic communication system (or even

[4]In accordance with what has become convention in linguistics (see e.g. Kirby 1999:9–13 for an overview), I will use the terms "functional," "non-functional" and "dysfunctional" as referring exclusively to the function of language in *communication* as opposed to any other function which certain properties of language may serve.

[5]They show, for example, that the relative arbitrariness of many features of a language merely constitutes a non-functional rather than dysfunctional property as long as the respective features are shared in a community of speakers.

a single symbol) to one as expressive and complex as language. This is the problem I will address in the following paragraphs, first theoretically and then by means of computer simulations.

5.2.1 Pragmatic plasticity as an "expressivity-enhancement tool"

The crucial difference between human language and the communication systems of other animals is not only (i) that the former is much more expressive than the latter, and (ii) that the former is learnt while the latter are innate, but also (iii) that the former is *creative* while the latter are not. Hockett (1960:8ff.) consequently mentions linguistic creativity, which he refers to as "productivity,"[6] alongside displacement, duality of patterning and traditional transmission (i.e. the fact that language is learnt rather than innate) as one of the defining design features of human language. In short, linguistic creativity describes the processes by which meanings which have never been expressed before come to be expressed. It is, of course, closely related to the emergence puzzle discussed in chapter 3. Linguistic creativity provides an explanation for why language is so highly adaptive to the ever-changing communicative needs of human societies: humans have a way of expressing novel meanings. And, as I will show in section 5.2.2 below, it thus also accounts for how a small symbolic communication system can grow bigger and become more expressive.

At some point, most computational models of language evolution face the problem of incorporating a process of linguistic creativity by means of which new meanings can be expressed. Hurford (2002) compares a set of computational models and points out that, in situations where their agents face the task of having to express a meaning for which they have no form available in their I-language, they resort to a process of invention by which they create a random new form for that meaning. This is not to say that the respective models make any theoretical claim about the role (or even presence) of invention in real language use. Supposedly, invention is merely used as a maximally abstract representation of other processes of linguistic creativity that are not further specified. Models that employ invention thus resort to somewhat of a deus-ex-machina

[6] So-called combinatorial productivity, "in the sense that one can coin new utterances by putting together pieces of familiar utterances, assembling them by patterns of arrangement also familiar in old utterances" (Hockett 1960:6), is often singled out from linguistic creativity in general. In this context, I refer to sections 3.3.1 and 3.4.3, where I have described how combinatorial productivity can be seen as a special case of pragmatic plasticity.

solution, which may be justifiable depending on the aim and focus of the respective model. Real invention, however, is relatively rare in language use; Trask (2000:369), for instance, only mentions a handful of examples, such as *blurb*, *quark* and *googol* and trademarks like *Teflon* and *Kodak*. Invention also presupposes that the innovating individual possesses some meta-linguistic knowledge about the nature of symbols. Having to assume such meta-linguistic knowledge is, of course, not ideal for an account of language evolution that aims to be rooted in pre-linguistic (i.e. pre-symbolic) cognitive capacities. As part of an explanation of how language became more expressive, invention can thus at best play a marginal role: while we cannot exclude the possibility that cases of invention have happened, it hardly seems to be the main process of linguistic creativity that was at work in the evolution of linguistic communication.

In combination with the process of conventionalisation, pragmatic plasticity constitutes an alternative form of linguistic creativity that does not have the problems that accompany the concept of invention. Pragmatic plasticity provides a method of expressing a new meaning by means of an extant symbol in contexts where it can be inferred that the conventional meaning of the produced form only serves as a clue that leads to the actually intended meaning. The mechanisms involved in this process, namely that conventional meanings can under- and/or overspecify the meanings they communicate, have been introduced in chapter 3, and how these mechanisms can be implemented in a computational model has been demonstrated in chapter 4. Pragmatic plasticity uses extant symbols as *stepping stones* to reach new meanings that could not have been expressed otherwise. When these new meanings become conventionalised, as explained in section 3.2.1, they can themselves serve as stepping stones to reach yet another set of previously inaccessible meanings. This process of iterated under- and/or overspecified use and subsequent conventionalisation can bring about the "ratchet effect" that Tomasello et al. (1993) describe for cultural evolution in general (see chapter 2). Pragmatic plasticity thus allows the users of a symbolic communication system to cumulatively explore new meaning spaces. I argue that it is like this that an initially small system of symbols can expand its expressivity and ultimately grow to the size of present-day human languages.

If we compare the two discussed forms of linguistic creativity, invention and pragmatic plasticity, we find that pragmatic plasticity has a greater explanatory potential and psychological plausibility with regard to language evolution. First, it is more frequent in present-day language use. Second, as I have argued in chapter 3, pragmatic plasticity is not specific to symbolic communication but constitutes a much older and more basic property that pertains to any form of ostensive-inferential communication. It does not presuppose the existence of symbolism nor the knowledge of what symbols are and how they work. Third, because of the way pragmatic plasticity works as an integral part of ostensive-inferential communication and because it only gradually turns into a new convention, it is a less conscious (but also slower) process of linguistic creativity than the purpose-driven invention of new symbols.

An issue that needs to be mentioned once more at this point (cf. section 4.1.2.2) is that of communicative success. The computational models discussed in Hurford (2002) employ invention to deal with cases where their agents would otherwise not be able to express a desired meaning. Pragmatic plasticity may reduce the number of such situations, but unlike the deus-ex-machina solution of invention, it cannot totally prevent the occurrence of conditions in which an agent does not have the means to express what he wants to say. However, I do not think that the fact that communication is not always feasible constitutes a problem in the first place. One can hypothesise, though, that the number of instances in which such conditions occur decreases as the expressivity of the developed symbolic communication system increases. This is a claim to be tested in the simulations to be introduced below.

In summary, I propose that pragmatic plasticity provides humans with a "tool" to enhance the expressivity of their symbolic communication systems. The introduced conceptual model of the cultural evolution of language suggests that the same processes that make ostensive-inferential communication possible and lead to the emergence of symbolic conventions also allow such systems to adapt their expressive power to the ever-changing communicative needs of human societies. I now turn to exploring this hypothesis by means of computer simulations.

5.2.2 *The development of expressivity in the model*

In chapter 4, I introduced a computational model that can be used to study the impact of pragmatic plasticity on the cultural evolution of language. By running various simulations, I will now investigate how expressivity evolves in the proposed model. I will apply the model in two main fashions. First, I will test the hypothesis that through iterated under- and/or overspecified use, the expressivity of a symbolic communication system continuously adapts to the communicative needs of its user. Second, I will employ the model as a laboratory-like environment to explore by what factors and how the development of the expressivity of a symbolic communication system is influenced. I define expressivity as the number of meanings a grammar can express, with the individual meanings being weighted according to their probability of occurrence. A detailed account of how the expressivity of a grammar is measured in the simulations is given in appendix B.1.

5.2.2.1 Behaviour under idealised conditions

I begin my investigation by describing the general behaviour of the model with respect to expressivity under simple, idealised conditions. Afterwards, I will discuss how that behaviour changes if individual parameters are altered. The setup I use as a starting point is the following. I assume that the agent can discern four semantic units A, B, X, Y, and that speaker meanings can contain up to two such units. This means that the semantic space in which he operates consists of the meanings A, B, X, Y, AB, AX, AY, BX, BY, XY.[7] The agent is given the capacity to produce two signals, X and Y, which are stored as self-symbolisations ($X \rightarrow X$ and $Y \rightarrow Y$) in his initial "grammar." The available phonological space thus consists of X, Y and XY. In each generation, a maximum of one semantic unit will be available as either inferable or ignorable information (refer to section 4.1.2.1 for the definitions of these two types of contextual information).[8] In this basic setup, I will work with the idealisation that once-acquired constructions are not lost even if their entrenchment has decayed considerably: the entrenchment threshold t is set to 0. In the case of synonymy, the agent chooses a signal at random: no specific synonym-selection strategy is assumed. The described conditions are summarised in (1).

[7] Note that meanings are *sets* of semantic units: the order in which the units occur is thus not relevant.

[8] I will use the term "context size" to refer to the amount of inferable and ignorable information that is available in an iteration.

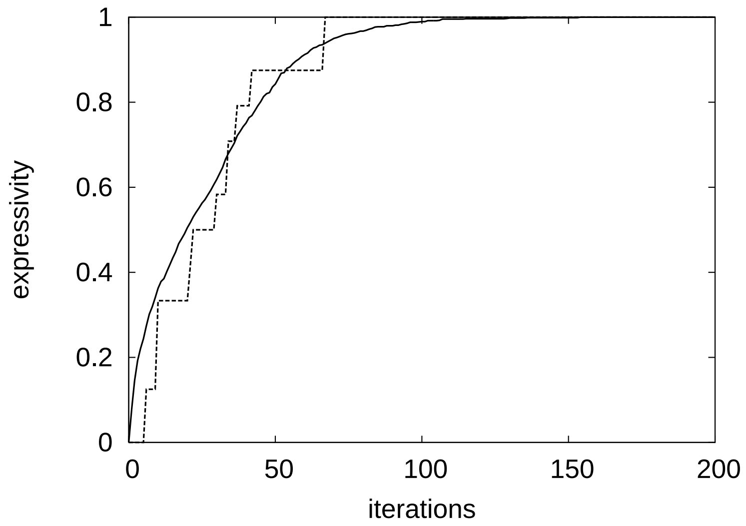


Figure 5.1: The development of the expressivity of an agent's grammar under idealised conditions. The solid line represents the average expressivity over 100 independent runs. The dashed line represents an individual run. The x-axis stands for the number of iterations the agent has been exposed to and the y-axis for the expressivity of his grammar. The parameter set underlying the simulation runs is specified in (1).

(1)	Semantic units:	*A*, *B*, *X*, *Y*
	Maximal meaning complexity:	2 semantic units
	Initial grammar:	$X \rightarrow X$
		$Y \rightarrow Y$
	Maximal context size:	1 semantic unit
	Entrenchment threshold:	$t = 0.0$
	Selection strategy:	random

Fig. 5.1 illustrates the way the expressivity of the agent's grammar develops under such simple, idealised conditions. It presents expressivity as a function of time or rather, more specifically, the number of iterations the agent has been exposed to. The first observation that can be made from Fig. 5.1 is that the expressivity of the agent's grammar rises steadily. There is no point in its development where expressivity decreases. This illustrates the cumulative effect of cultural evolution: if nothing is lost, the "ratchet" does not slip back. The second observation that can be made is that the agent's grammar eventually achieves maximal expressivity. It reaches a state where every possible speaker meaning

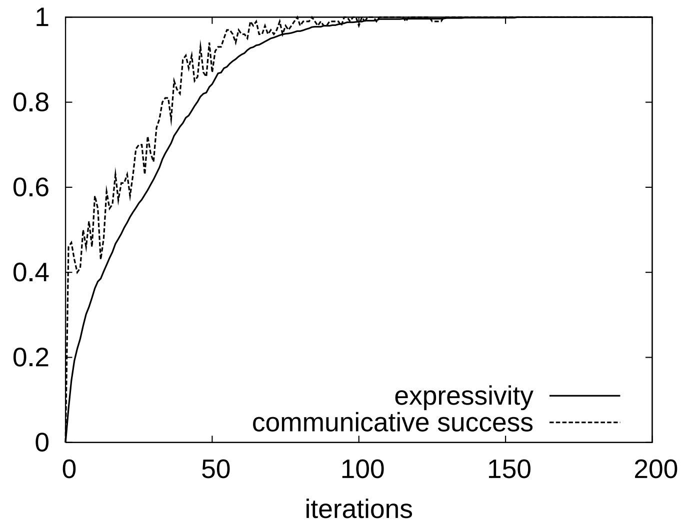


Figure 5.2: The development of communicative success. Both lines represent the average of the respective value over 100 simulation runs under the conditions given in (1).

is represented as a conventional meaning in one of its constructions. In the presented setup, the expressivity of the agent's grammar thus rises continuously until it finally reaches its maximum.

Fig. 5.2 plots the average communicative success in each iteration (an iteration is successful if the agent is able to produce an appropriate signal for the intended speaker meaning) and compares the development of this number over the course of the simulation with the development of the expressivity of the agent's grammar.[9] This figure not only shows how much of the semantic space the agent's grammar covers but also how successful that grammar is in addressing the agent's communicative needs. It can be observed that initially, the agent is less likely to encounter an iteration where he is able to communicate the given speaker meaning. However, as the expressivity of his grammar grows, this probability increases, until eventually, the agent can deal with every communicative

[9]Note that the term "communicative success" is used in a purely speaker-oriented sense here: the agent's communicative act is considered successful if he is able to produce a signal that conveys the given speaker meaning in the given context according to the procedures defined in the production algorithm detailed in appendix A. The hearer's role in successful communication is only implied; the model rather makes the idealising assumption that a signal that meets the requirements of the given context will successfully convey the intended speaker meaning.

situation he is exposed to. We find here again the fact that the conventionalisation of form-meaning associations leads to increased context-independence. As long as the agent does not possess a conventionalised construction for every possible meaning, the success of his communicative efforts depends on the specifics of the contexts in which he attempts to convey his message. Once the agent's grammar has reached maximal expressivity, that context does not have an impact anymore on whether the agent will be able to communicate the intended speaker meaning or not. While he still may employ under- and/or overspecification, he can always resort to fully specifying the intended meaning if necessary.

What Figs. 5.1 and 5.2 show is how the expressivity of a symbolic communication system adapts to a *static* environment. I work with a static environment, that is, with the idealised assumption that the communicative needs of the agent remain the same over time, because it allows us to study the adaptive behaviour of the system more clearly: a communication system is the better adapted to the communicative needs of its user the more expressive it is and the less frequently the user is incapable of conveying an intended message. We have seen that under the described idealised conditions, the communication system fully adapts to its environment through iterated under- and/or overspecified use. However, the real environment in which human language evolves is *dynamic*: human communicative needs change over time as new concepts need to be expressed and old concepts are not referred to anymore. We can infer that a communication system would maintain the observed general behaviour, that is, that it would continuously move toward maximising its adaptation to the communicative needs of its user, even in circumstances where the environment in which it unfolded were dynamic, but that because the communicative needs would keep changing, it would never quite reach maximal expressivity or would not maintain that state for a long time.

5.2.2.2 Influencing factors

I will now investigate how the development of the expressivity of the agent's grammar changes if individual parameters are modified. To this end, I will take the basic setup given in (1), modify its parameters and compare the behaviour of the model under the modified conditions with that under the original conditions. The factors I will look at include the size of the semantic and the phonological space, the complexity of the speaker meaning, the amount of inferable and ignorable information available in the context, the degree of loss, the employed

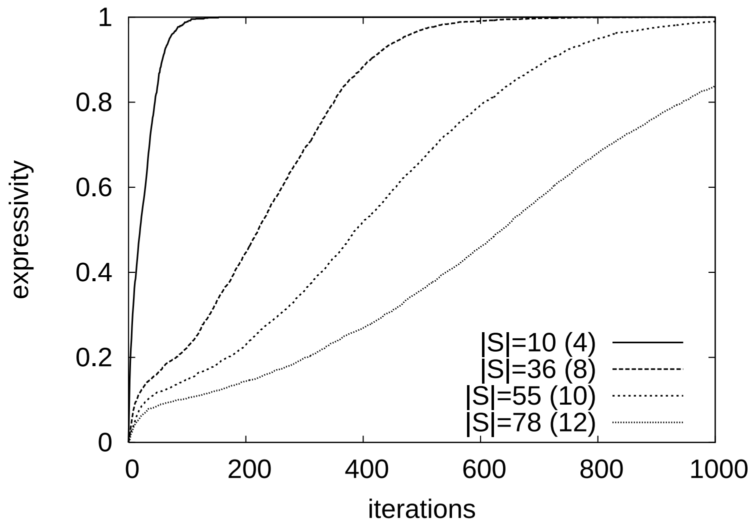


Figure 5.3: The development of expressivity with semantic spaces of different size. Each line represents the average development of expressivity over 100 independent simulation runs. $|S|$ indicates the size of the semantic space; the numbers in parentheses indicate the number of elementary semantic units that occur in the semantic space. All other parameters are as in setup (1).

selection strategy, and the structure of the semantic space. The aim is to explore if certain conditions facilitate the adaptation of a symbolic communication system to the communicative needs of its user, that is, if expressivity rises more quickly and reaches higher levels under some conditions than under others.

The size of the semantic space. In the original setup, the semantic space is defined by the four semantic units *A, B, X, Y* and the fact that individual speaker meanings consist of maximally two such semantic units. I now describe what happens if one increases the number of identifiable semantic units while retaining the meaning complexity. The conditions listed in (2), for instance, are identical to those in (1) with the exception that twice as many semantic units can be identified by the agent: instead of the original four semantic units *A, B, X, Y*, the agent can now discern eight, namely *A, B, C, D, E, F, X, Y*, which increases the size of the semantic space from 10 to 36 possible speaker meanings.

(2)	Semantic units:	*A, B, C, D, E, F, X, Y*
	Maximal meaning complexity:	2 semantic units
	Initial grammar:	$X \rightarrow X$
		$Y \rightarrow Y$
	Maximal context size:	1 semantic unit
	Entrenchment threshold:	$t = 0.0$
	Selection strategy: :	random

Fig. 5.3 illustrates the development of the expressivity of the agent's grammar relative to the size of the meaning space. It shows that, other conditions being equal, a symbolic communication system adapts more quickly to a smaller meaning space than to a larger one: where the semantic space contains only 10 meanings, maximal expressivity is achieved within 100 successful iterations, whereas with a semantic space of size 36, this state is only reached after 500 successful iterations; and it takes even longer for semantic spaces of size 55 and 78. Because the agent's grammar expands cumulatively over time, a smaller semantic space is covered more quickly than a larger one if the starting point is the same.

The size of the phonological space. A higher number of semantic units also provides for more possibilities to modify the phonological space. In (3) below, I have taken the conditions given in (2) but extended the agent's initial grammar: he is now capable of producing not only *X* and *Y* but also *W* and *Z*. The subregion of the semantic space that constitutes the phonological space has therefore been extended from 3 possible meanings (*X, Y, XY*) to 10 (*W, X, Y, Z, WX, WY, WZ, XY, XZ, YZ*).

(3)	Semantic units:	*A, B, C, D, W, X, Y, Z*
	Maximal meaning complexity:	2 semantic units
	Initial grammar:	$W \rightarrow W$
		$X \rightarrow X$
		$Y \rightarrow Y$
		$Z \rightarrow Z$
	Maximal context size:	1 semantic unit
	Entrenchment threshold:	$t = 0.0$
	Selection strategy: :	random

Fig. 5.4 illustrates what impact the size of the phonological space has on the development of the expressivity of the agent's grammar. It can be observed that a

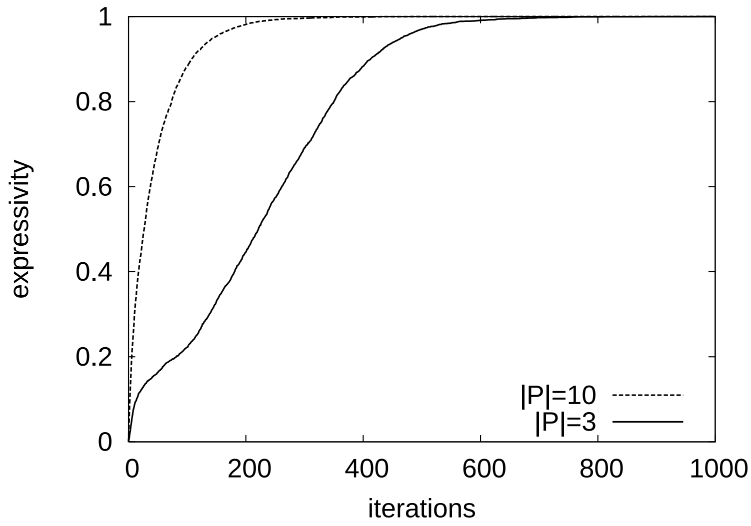


Figure 5.4: Development of expressivity with phonological spaces of different size. Either line represents the average development of expressivity over 100 independent simulation runs. The dashed line depicts the situation where the size of the phonological space P is 10, the solid line where it is only 3. All other parameters are as in setup (2).

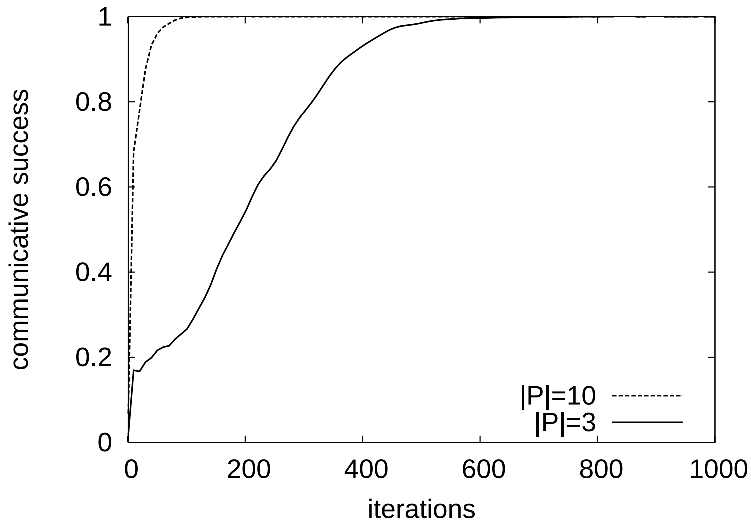


Figure 5.5: Development of communicative success with phonological spaces of different size. The depicted conditions are the same as in Fig. 5.4. Either line represents the average communicative success in a particular iteration over 100 independent simulation runs.

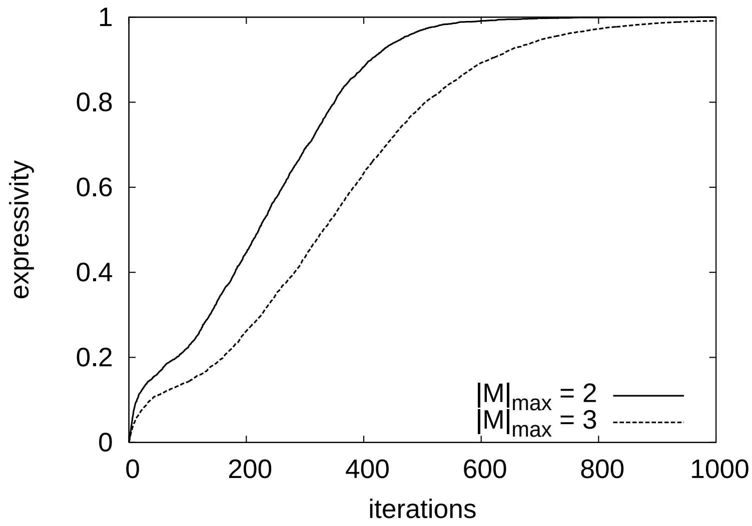


Figure 5.6: Development of expressivity with different maximal meaning complexities. The solid line plots the average development of expressivity over 100 independent simulation runs with maximal meaning complexity $|M|_{max}$=2 (i.e. a speaker meaning consists of maximally two semantic units). The dashed line represents the average development of expressivity over 100 independent runs with maximal meaning complexity $|M|_{max}$=3. In both cases, all other conditions are as in (2).

larger phonological space gives the agent a "head start": because more meanings can be expressed independently of the specific context by producing direct evidence for them, the agent's communication system is initially more successful at conveying the intended messages, and its expressivity consequently grows more rapidly during that phase. This explanation for the difference between the two cases depicted in Fig. 5.4 is confirmed by Fig. 5.5, which shows the development of communicative success under the same conditions. The figure demonstrates that a larger phonological space, or rather a larger initial grammar, means that more iterations are successful right from the beginning, whereas with a smaller initial grammar, iterations are only successful very rarely as long as not many form-meaning associations have become conventionalised yet.

Meaning complexity. The size of the semantic space can be modulated not only by the number of semantic units that an agent is capable of discerning but also by the maximal complexity assumed for meanings. In the simulations discussed so far, this complexity was set to 2: speaker meanings consisted of up to two

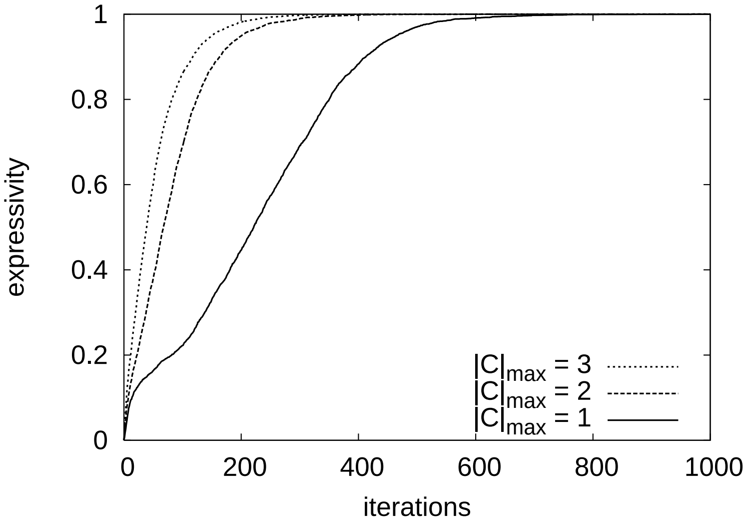


Figure 5.7: Development of expressivity with different context sizes. The dotted line plots the average development of expressivity over 100 independent simulation runs with maximal context size $|C|_{max}$=3, the dashed line represents the same for maximal context size $|C|_{max}$=2, and the solid line for maximal context size $|C|_{max}$=1. All other conditions are as in (2).

semantic units. An increase of this number also means that a greater variety of meanings can occur, so that we have to expect an effect similar to that of an increase of the number of semantic units discussed above. This hypothesis is confirmed by Fig. 5.6, which illustrates the development of the expressivity of the agent's grammar relative to different maximal meaning complexities. As expected, it can be observed that, on average, expressivity rises slower and takes more time to reach its maximum if speaker meanings are more complex.

The amount of inferable and ignorable information (context size). Fig. 5.7 illustrates the effect that an increased context size has on the development of the expressivity of the agent's grammar: the more contextual information is available, the more quickly expressivity rises on average. This makes sense from a theoretical perspective too: the more material can be used to under- and overspecify the intended speaker meaning, the larger will the degree to which pragmatic plasticity is employed be. And since it is through the use of pragmatic plasticity that new form-meaning associations become conventionalised, more possibilities for pragmatic plasticity would have to lead to a faster increase in the expressivity of the respective symbolic communication system. At the same time, the agent

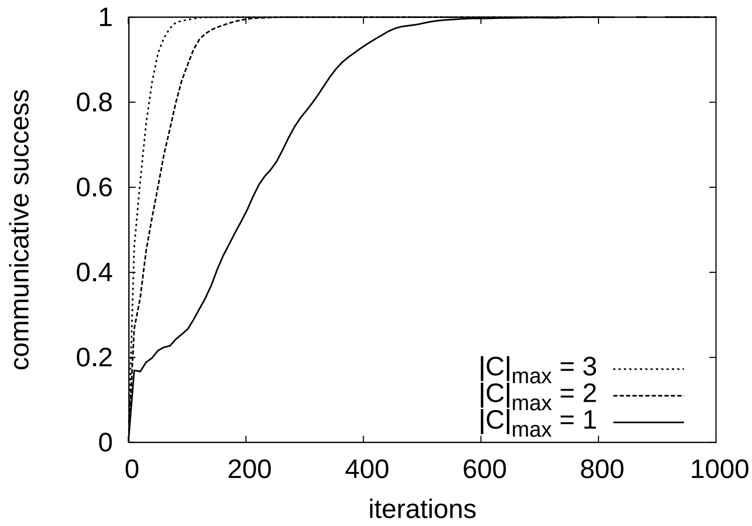


Figure 5.8: Development of communicative success with different context sizes. The depicted conditions are the same as in Fig. 5.7. Either line represents the average communicative success in a particular iteration over 100 independent simulation runs.

encounters fewer iterations in which he is not capable of producing an appropriate signal if he can make use of under- and/or overspecification more often because, on average, larger amounts of inferable and ignorable information are available. This observation is illustrated in Fig. 5.8, which plots the development of communicative success for the three different maximal context sizes.

In summary, it can be observed that an increase in the complexity of speaker meanings and an increase in the amount of inferable and ignorable information that is available in an iteration have opposite effects. The more sophisticated the meanings that the agent needs to express are, the more slowly the expressivity of his communication system rises; but the more contextual information he can make use of in individual communicative acts, the faster the expressivity of his communication system grows. This observation is summarised in Fig. 5.9. Increasing the maximal context size raises the curve that represents the development of expressivity; increasing the maximal meaning complexity lowers it. An increased maximal context size can thus cancel the effect of an increased maximal meaning complexity, and vice versa. This is the important insight here: the effect of an individual factor, as clear as it may appear in isolation, can be mitigated by

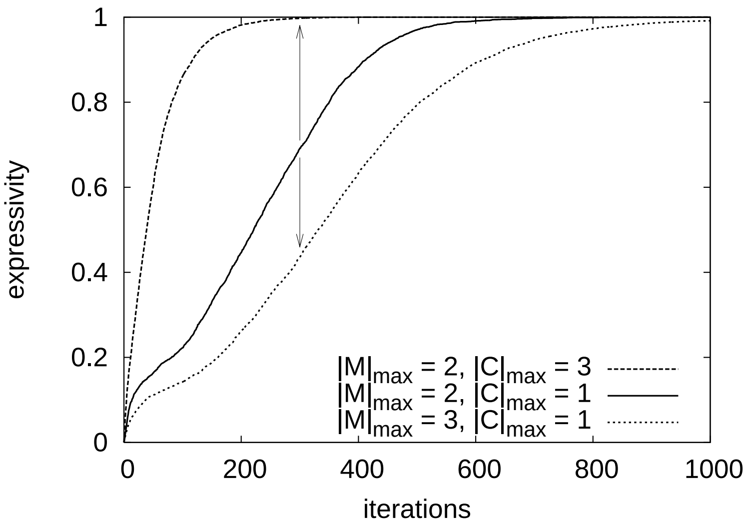
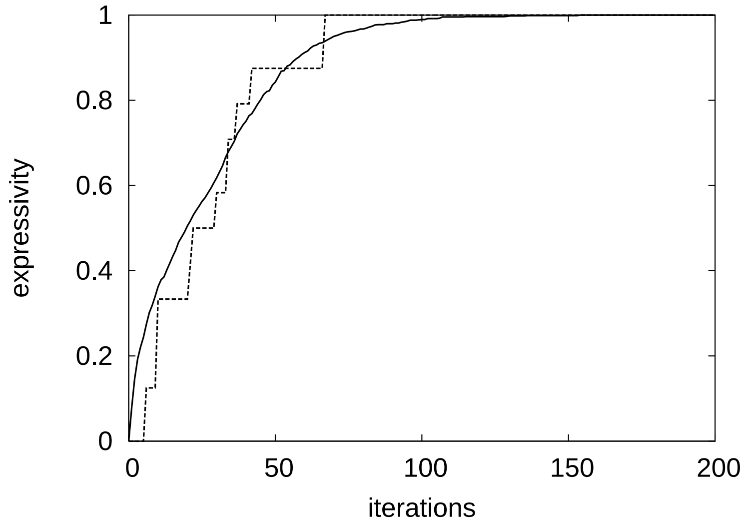


Figure 5.9: Comparison of the opposite effects of meaning size and context size on expressivity. Each line plots the average development of expressivity over 100 independent simulation runs under the conditions given in (2). The solid line represents a "neutral" case with maximal meaning complexity $|M|_{max}$=2 and the maximal context size $|C|_{max}$=1. In comparison to the solid line, the dashed line shows the effect of increasing the maximal context size (from 1 to 3), and the dotted line shows the effect of increasing the maximal meaning complexity (from 2 to 3).

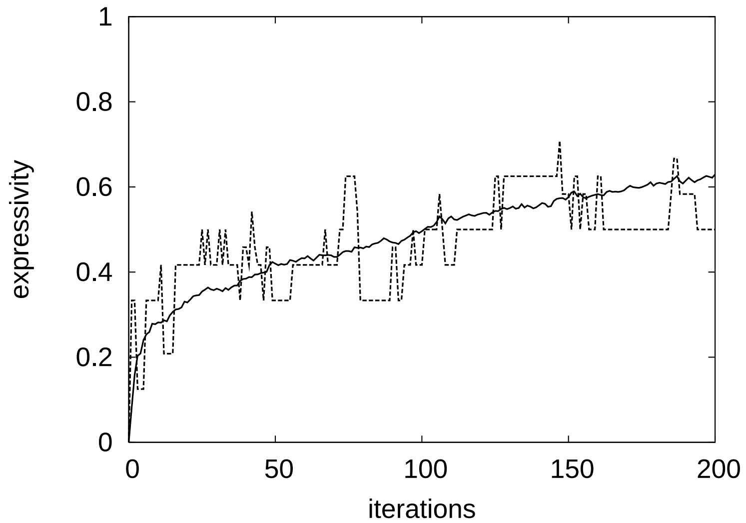
the effect of another factor that occurs together with it. A large semantic space may take long to explore if only little contextual information is made use of but can be covered quickly if large amounts of inferable and ignorable information are available.

The rate of loss. The development of the expressivity of an agent's grammar changes if loss is introduced into the simulation. Remember that the computational model provides the option to define an entrenchment threshold: any form-meaning association whose entrenchment falls below that threshold will be lost (i.e. "forgotten") by the agent (cf. section 4.1.3.2). The threshold can be set to any value between 0 and 1. If the threshold is set to 0, as it was in all simulation runs discussed so far, no form-meaning pair is ever lost; if it is set to 1, none is ever remembered. The entrenchment of a construction is enforced through use and decays through lack of use.

Figs. 5.10(a) and 5.10(b) illustrate the effect that the introduction of loss has on the development of the expressivity of the agent's grammar. Both figures show



(a) without loss ($t = 0.0$)



(b) with loss ($t = 0.5$)

Figure 5.10: Development of expressivity with and without loss. In both figures, the solid line represents the average over 100 independent simulation runs and the dashed line represents an individual run. (a) The development of the expressivity of the agent's grammar without loss. (b) The development of the expressivity of the agent's grammar with loss: any form-meaning mapping whose entrenchment fell below the threshold $t = 0.5$ was removed from the agent's grammar. All other parameters are as specified in setup (1).

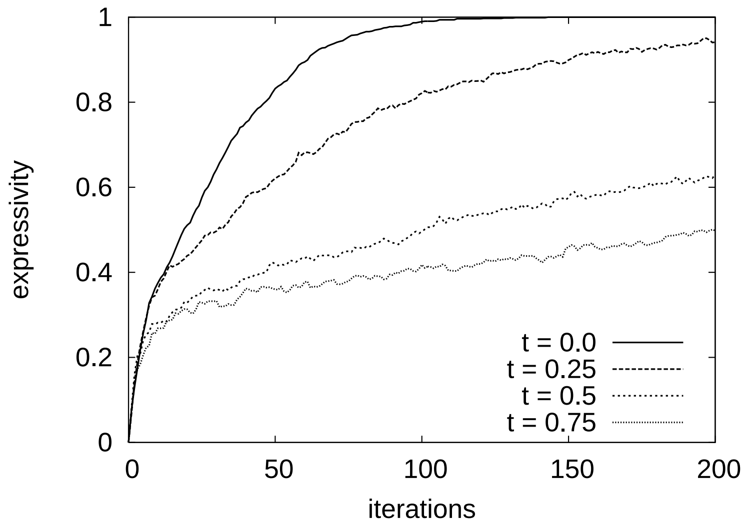
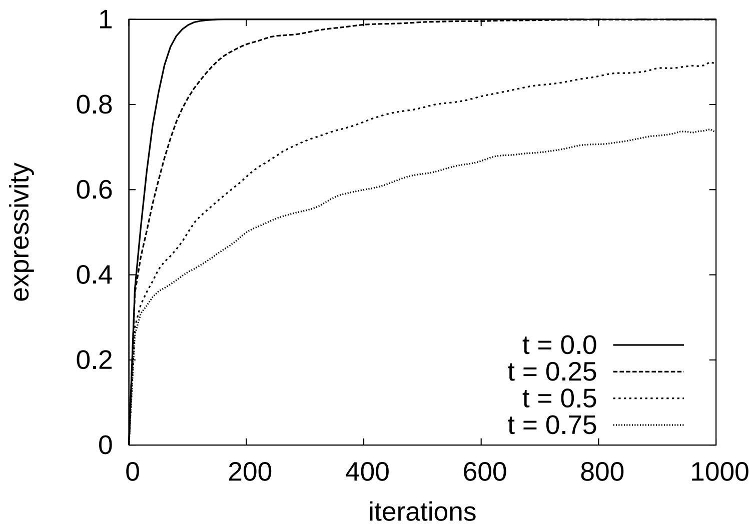


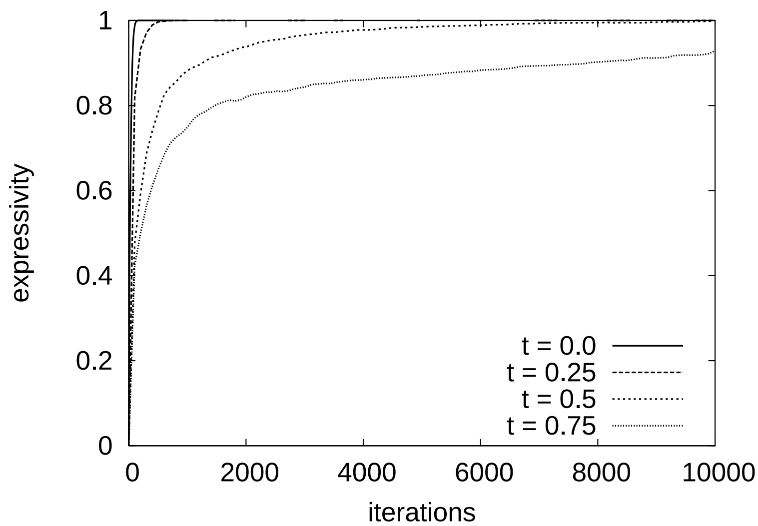
Figure 5.11: Development of expressivity with different loss rates. Each line represents the average over 100 independent simulation runs. The degree of loss (determined by the entrenchment threshold *t*) varies between the lines; all other parameters are those listed in (1).

simulations run under the conditions given in (1) above, with the only difference that Fig. 5.10(a) renders the situation as it is if there is no loss whereas Fig. 5.10(b) plots the development if there is loss. A first observation to be made is that if loss is introduced to a simulation, the expressivity of the agent's grammar does not rise steadily anymore: it experiences frequent setbacks. While in Fig. 5.10(a), expressivity increases continuously, in Fig. 5.10(b), it often falls back to lower levels again—even though an overall rise of expressivity can still be observed. This picture complies with the claims made by the ratchet model of cultural evolution (cf. section 2.1.2.2): for the ratchet effect to be most effective, cultural artefacts have to be preserved as faithfully as possible over time; the more loss occurs, the more often the ratchet "slips back."

A second observation that can be made is that, as a consequence of the slippage introduced by loss, expressivity rises more slowly if loss occurs. This fact is documented in Fig. 5.11, which shows the development of the expressivity of the agent's grammar for different rates of loss. The general rule that becomes evident from this figure is that the higher the rate of loss, the more slowly expressivity increases. The question arises whether expressivity also stabilises at lower levels



(a)



(b)

Figure 5.12: Development of expressivity with different loss rates over longer time. These two figures plot the average development of the expressivity of the agent's grammar under the conditions shown in Fig. 5.11 over the course of (a) 1,000 iterations and (b) 10,000 iterations.

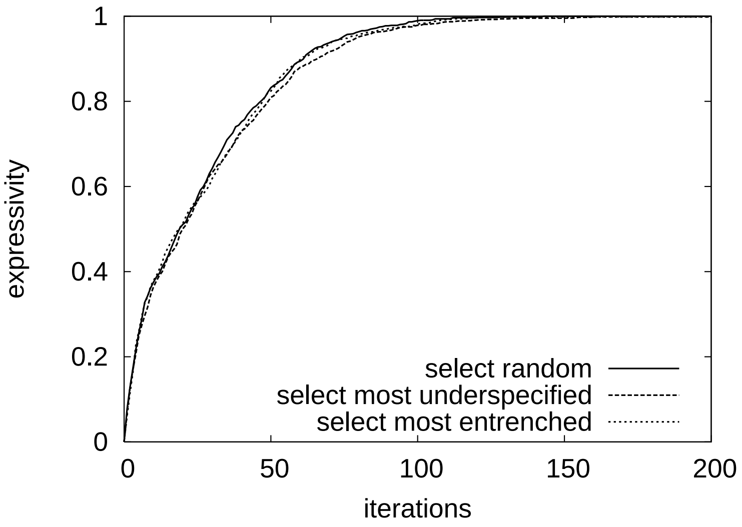
and does not reach its maximum if the agent's grammar is exposed to a relatively high degree of loss. Fig. 5.11 does not suffice to answer this question: while the expressivity curves for the cases with loss do not reach the theoretical maximum within the represented number of transmissions and flatten much earlier than the curve of the case without loss, they continue to rise—if ever so slightly—till the end of the timespan shown. Fig. 5.12(a) represents the same set of simulations but plots the development of expressivity over not just 100 but 1,000 iterations. We recognise that all curves exhibit much higher levels of expressivity: one of them even reaches the maximum, and none of them has completely flattened by the end of 1,000 iterations. This trend is repeated in Fig. 5.12(b), which represents the development over 10,000 iterations. Within this timespan, another curve has all but reached the maximum ($t = 0.5$), and even the grammars with the highest represented entrenchment threshold ($t = 0.75$) have reached an expressivity of more than three quarters. We can conclude that even though grammars that are exposed to higher rates of loss grow much more slowly, they still seem to come to exhibit high expressivity after a sufficient amount of time.

The simulations provide us with two interesting insights about the evolution of the expressivity of a symbolic communication system. First, memory is important. The more past usage events are remembered, and thus available for re-use in future usage situations, and the longer they are remembered, the more likely a system is to get "off the ground" within a relatively short timespan. Second, time and the frequency with which communicative interactions happen are also important. Reduced memory does not necessarily mean that a symbolic communication system does not evolve at all—it simply takes much more time, or rather a much higher number of communicative interactions. Even with a relatively small percentage of past usage events being remembered, such a system can eventually become as expressive as one that grew in conditions where everything is remembered. These insights allow us to speculate about some of the reasons why humans have evolved language but other apes have not: an increased memory and a drive (or need) for more frequent communicative interaction may be part of what made the difference.

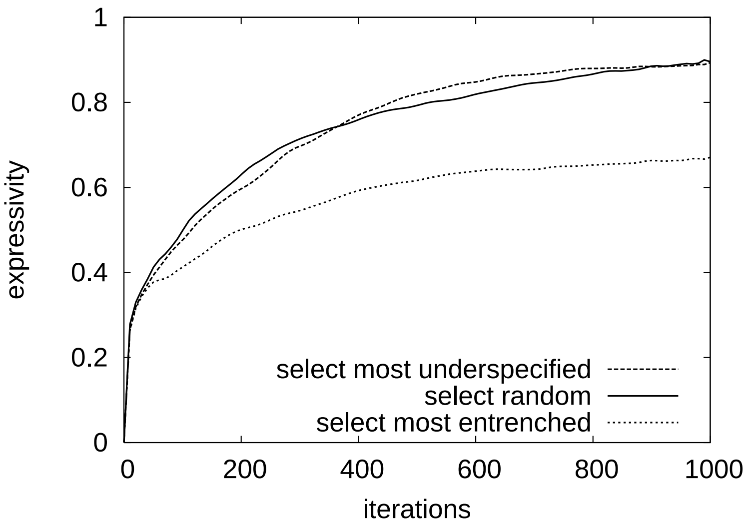
The synonym-selection strategy In section 4.1.2.3, I have explained that the agent sometimes faces situations in which more than one signal meets the requirement of being able to convey the intended speaker meaning. In such situations, the agent has to choose one of the possible signals. I have also pointed out that humans select signals for a variety of reasons—many of them sociolinguistic or register-related—and that therefore the neutral way of modelling this process is to have the agent choose one of the possible signals at random. On the other hand, pragmatic theories suggest that such sociolinguistic reasons only come to work *after* a number of more fundamental pragmatic principles have been applied. In the model, these principles correspond to the agent choosing either (i) the signal whose signal meaning underspecifies the speaker meaning most or (ii) the signal where the constructions involved in producing the signal are most entrenched. I will therefore compare simulation runs in which one of these two strategies has been applied consistently by the agent with simulation runs in which the agent has chosen randomly.

The first observation to be made is that if there is no loss, the employed selection strategy does not have an impact on the development of the expressivity of the agent's grammar. This is illustrated by Fig. 5.13(a). The figure shows no variation in the behaviour of the model: no matter what selection strategy the agent applies, the expressivity of his grammar evolves in the same way. The picture changes once loss is introduced into the simulations. Like Fig. 5.13(a) before, Fig. 5.13(b) plots the development of the expressivity of the agent's grammar under the employment of various selection strategies—but now under the condition that there is a certain degree of loss. A comparison of the two figures shows that the presence of loss entails that the choice of a selection strategy has an impact on the development of expressivity. In general, we can observe an increased variation between the plotted lines.

The most conspicuous effect can be observed if the agent consistently chooses the most entrenched signal. This strategy performs worse than the other two: the expressivity of the agent's grammar increases slower under this condition. At first, this finding may seem surprising. It can, however, be explained. By producing a signal, an agent also further entrenches the form-meaning mappings involved in the production of that signal. If in any situation where more than one signal is possible, the signal that is most entrenched is selected, the differences between the entrenchment of individual mappings become more prominent. Over time, the result is a grammar that contains a number of highly entrenched constructions and a number of scarcely entrenched constructions. Whenever a case of synonymy occurs, this dichotomy is further enforced: some already highly entrenched form-meaning associations are entrenched further and the only weakly entrenched ones experience further decay. The result of this dynamic is that more



(a) without loss ($t = 0.0$)



(b) with loss ($t = 0.5$)

Figure 5.13: Development of expressivity with different selection strategies being employed. Both figures plot the development of the expressivity of the agent's grammar if specific strategies are employed to select among synonyms. Each line represents the average over 100 simulation runs. The assumed conditions are those listed in (2). (a) Without loss (entrenchment threshold $t = 0.0$): the choice of selection strategy has no impact on the development of expressivity. (b) With loss (entrenchment threshold $t = 0.5$): the impact of the chosen selection strategy shows more variation. The strategy of selecting the most entrenched signal performs worse than the other strategies because it amplifies the negative effect that loss has on the growth of expressivity.

loss occurs because there are more constructions with only minimal entrenchment. And, as we have already seen above, a higher degree of loss means that the expressivity of the agent's grammar cannot rise as fast as it would otherwise. The psychologically most plausible selection strategy (cf. section 4.1.2.3), namely to choose the most entrenched signal, thus performs worst if exposed to loss because it *amplifies* the effect that loss has on the evolution of a grammar's expressivity.

The structure of the semantic space. Cross-linguistic studies, in particular of grammaticalisation, show that there is a certain regularity to semantic change (e.g. Heine and Kuteva 2002; Traugott and Dasher 2005). One consequence of this is the relative uni-directionality of grammaticalisation (see e.g. Hopper and Traugott 2003:ch. 5). As I have shown in chapter 3, what becomes conventionalised, and therefore how a linguistic code changes semantically, crucially depends on the specific types of inferences that can be drawn in individual usage situations. A large part of how the English expression *be going to* became grammaticalised, for instance, depends on the fact that an act of spatial motion frequently involves intention. Here we have a case of two concepts co-occurring in the world in a way that the one can often be inferred from the presence of the other. Semantic space is thus not totally random but exhibits some structure: in general, certain concepts are communicated by agents more frequently than others, and some concepts often occur in the company of others.

Since such environmental structure appears to play a crucial role in many known phenomena of semantic change, it makes sense to look at the impact that patterns in the semantic space may have on the development of the expressivity of the agent's grammar. Up to this point, the semantic space of the simulations run had a *flat* structure: all semantic units were equally likely to occur in speaker meanings, and all semantic units were equally likely to occur as part of the contextual information. I will now compare the behaviour of the model under such circumstances with its behaviour in cases where its semantic space does not have a flat structure, that is, with cases where not all semantic units occur with equal probability.

I hypothesise that especially two aspects of the probability structure of a semantic space have an impact on how the expressivity of an emerging communication system develops: (i) the likelihood for producible semantic units (i.e. semantic units that are part of the agent's phonological space and can thus directly be

Semantic unit	Converse structure		Parallel structure		Frequency of producible and non-producible units
	p_M	p_C	p_M	p_C	
	Pattern 1		**Pattern 2**		
A	0.49	0.01	0.49	0.49	*Non-producible units frequent*
B	0.49	0.01	0.49	0.49	
X	0.01	0.49	0.01	0.01	*Producible units infrequent*
Y	0.01	0.49	0.01	0.01	
	Pattern 3		**Pattern 4**		
A	0.01	0.49	0.01	0.01	*Non-producible units infrequent*
B	0.01	0.49	0.01	0.01	
X	0.49	0.01	0.49	0.49	*Producible units frequent*
Y	0.49	0.01	0.49	0.49	

Table 5.1: Four probability patterns for the semantic space of the basic setup. The semantic units X and Y are directly producible (i.e. they are part of the agent's phonological space), the semantic units A and B are not directly producible.

produced as units of form) to occur in the speaker meaning, and (ii) the likelihood of non-producible semantic units to be inferable from context. The intuitive assumption would be that a system "kicks off" more easily if the available producible units frequently occur in the meaning the agent needs to communicate. Likewise, we can hypothesise that the expressivity of a grammar is more likely to expand substantially if non-producible meanings can frequently be inferred from the context.

To study these two hypotheses, I have devised the four different probability patterns listed in Table 5.1. Each pattern represents the probability distributions of a semantic space that is made up of four elementary semantic units (A, B, X, Y) as in setup (1). Two of these units (X and Y) can directly be produced by the agent—they constitute units of form; the other two units (A and B) are not directly producible. Every unit is assigned two probability values: (i) the probability with which it occurs as part of the speaker meaning that the agent needs to communicate in an iteration, and (ii) the probability with which it occurs as

Value for P_M	Value for P_C	Effect
high	high	likely to occur in the speaker meaning and likely to be inferable
high	low	likely to occur in the speaker meaning but unlikely to be inferable
low	high	unlikely to occur in the speaker meaning but likely to be ignorable
low	low	unlikely to occur in the speaker meaning but also unlikely to be ignorable

Table 5.2: The effect of specific probability combinations on the occurrence of individual semantic units.

part of the contextual information, that is, the probability with which it constitutes inferable or ignorable information in a particular usage situation. I refer to the former as P_M and to the latter as P_C. The distributions of P_M and P_C do not necessarily need to stand in any relation to each other. However, to study the effects of the structure of the semantic space, I will focus on two limiting cases that deviate most clearly from the flat structure applied so far. The first case, found in patterns 2 and 4, is one where the two probability distributions are parallel: P_M and P_C are the same for any semantic unit. In the second case, found in patterns 1 and 3, the two distributions are converse: P_M is high where P_C is low, and vice versa.

To understand what it means for a semantic unit to have high or low values for P_M and P_C, we need to have a brief look at the specifics of the algorithm that generates the contextual information (i.e. inferable and ignorable meaning) for each iteration.[10] This algorithm works as follows. It first randomly selects a number of semantic units according to their probabilities of occurrence. Then, the algorithm determines for each selected unit whether it is meant to be inferable or ignorable in the present context. Units that have also been selected to be part of the speaker meaning are turned into inferable meaning, units that do not at the same time occur in the speaker meaning constitute ignorable meaning. Table 5.2 lists the effects that the combination of specific values for P_M and P_C can have on

[10]The details of the algorithm are given in appendix A.3.

the occurrence of individual semantic units. A semantic unit with high values for both P_M and P_C is likely to occur in a speaker meaning and to be inferable at the same time: it can frequently be left underspecified. A semantic unit with a high value for P_M but a low value for P_C, on the other hand, is likely to occur in a speaker meaning but unlikely to be inferable: it will frequently have to be specified explicitly. If, on the other hand, a semantic unit has a low value for P_M, it is unlikely to occur in a speaker meaning—but if, at the same time, its value for P_C is high, it will be likely to constitute ignorable meaning and thus often be available for use in overspecification.

The four patterns in Table 5.1 are also different from one another with regard to whether the producible units X and Y frequently occur as part of the speaker meaning. In patterns 1 and 2, they only seldom constitute part of the meaning that the agent needs to communicate, whereas they are frequently part of the speaker meaning in patterns 3 and 4. In summary, the four patterns can be described as follows:

- *Pattern 1:* The speaker meaning frequently consists of semantic units that are not directly producible as forms. However, these units are rarely inferable from the context. Producible units, on the other hand, often constitute ignorable information but seldom occur as part of the meaning that the agent needs to communicate.
- *Pattern 2:* The speaker meaning frequently consists of semantic units that are not directly producible as forms. These units are also often inferable from the context. Producible semantic units rarely occur at all, whether in the speaker meaning or as inferable or ignorable information.
- *Pattern 3:* The speaker meaning frequently consists of units that are directly producible as forms. Non-producible units rarely occur in the speaker meaning, but if they do, they are likely to be inferable from context. They also often constitute ignorable information.
- *Pattern 4:* The speaker meaning frequently consists of directly producible units, which are often inferable from the context at the same time. Non-producible units, on the other hand, rarely occur at all, whether as part of the speaker meaning or as inferable or ignorable information.

Fig. 5.14 plots the development of the expressivity of the agent's grammar for the four patterns as well as for a semantic space with a flat probability distribution. It confirms the two formulated hypotheses. First, the system gets off the

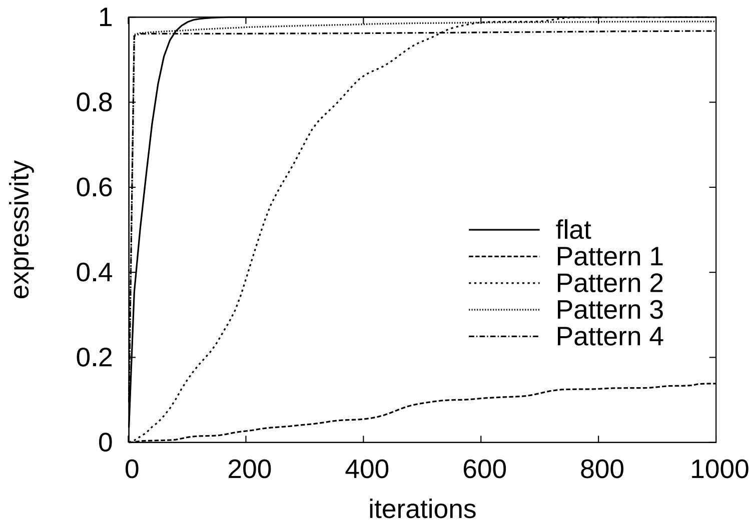


Figure 5.14: Development of expressivity for semantic spaces with different probability patterns. Each line plots the average development of the expressivity of the agent's grammar over 100 independent simulation runs. The conditions are as in setup (1); only the probability structure of the semantic space varies between the lines. The individual probability patterns are detailed in Table 5.1.

ground faster if the speaker meaning frequently consists of producible semantic units: the initial increase of expressivity is substantially steeper for patterns 3 and 4, where this is the case, than for the flat distribution. Where the occurrence of producible units in the speaker meaning is rare (patterns 1 and 2), the initial rise of expressivity is even slower. The second hypothesis stated that systems grow more expressive if non-producible units are frequently inferable from context. This hypothesis is confirmed in the development for patterns 2 and 3 compared to patterns 1 and 4 respectively. Pattern 1 cannot get its expressivity off the ground: producible units rarely occur in the speaker meaning and non-producible meanings are seldom inferable. Pattern 2, in contrast, also rarely encounters producible units as part of the speaker meaning. However, after a certain delay, it still manages to reach maximal expressivity because, once the groundwork has been laid, the frequently occurring non-producible meanings can be inferred from the context. Patterns 3 and 4 are both very successful initially because producible units often constitute part the speaker meaning. But, in contrast to pattern 3, pattern 4 fails to reach maximal expressivity (at least within the displayed 1,000 iterations) because the remaining non-producible meanings are only rarely inferable from the context. The analysed simulations thus suggest that the expressivity of a symbolic communication system grows the faster the more likely the communicated meanings are to contain semantic units that are directly producible as forms and the more frequently non-producible semantic units can be inferred from the context.

5.2.2.3 Summary and interpretation

With the described simulation runs I have shown that the developed conceptual and computational model of cultural evolution has the explanatory capacity to account for how, over time, the expressivity of a grammar adapts to the communicative needs of its users. The simulations have all been run in a static environment: the semantic space did not change over time. But since the mechanisms by which the grammar adapts to its environment remain the same whether or not that environment is static, we must assume that even in the fast-changing environment of human societies, language continuously adapts to whatever its environment looks like at the moment. Its adaptive processes are blind to the fact that the semantic space might have changed since the last usage event. The core point that the simulations illustrate, however, is how the expressivity of a grammar can grow from "nothing," and how the number of situations in which the agent is not capable of communicating what he wants to goes down as the expressivity of his grammar increases.

The simulation runs discussed here have also shown that the degree to which a symbolic communication system comes to be adapted to the communicative needs of its users, and the time it takes to get there, can be facilitated or hindered by a variety of factors, such as (i) the degree of loss, (ii) the size of the semantic space, (iii) the relative size of the phonological space, (iv) the complexities both of the available contextual information and of the speaker meanings to be communicated, (v) the employed strategy to choose from among synonyms, and (vi) the probability structure of the semantic space. Under the conditions I have looked at, the following factors have proven to be advantageous to the adaptive process: (i) a smaller semantic space, (ii) a larger phonological space, and, within the same semantic and phonological space, (iii) larger amounts of available contextual information, (iv) less complex speaker meanings, and (v) a minimised amount of loss. Especially in the initial phase of the evolution of a symbolic communication system it seems to be helpful if either producible semantic units occur frequently in the speaker meaning or otherwise non-producible semantic units are at least frequently inferable from the context.

It could also be observed that the impact of some factors only comes to bear if there is a certain degree of loss: which strategy the agent employs to choose from a set of synonyms, for instance, has no bearing on how well a communication system adapts as long as there is no loss, but if there is loss, its negative effect is even amplified if the agent always chooses to produce the most entrenched signal. The simulations illustrate that language evolution is a complex adaptive system. Some parameters in this system have a very clear effect (the absence of loss, for instance, is always advantageous) whereas the impact of others varies depending on what other parameters they co-occur with. This fact demands that any generalisation derived from the simulations must come with the caveat that, ultimately, all observations may be specific to the particular conditions under which they have been made.

5.3 Signal economy

Signal economy is another aspect of language that can convey the appearance of design for communication. Historical linguists have documented that the forms of individual constructions are often reduced in the course of language change. From this observation, they infer that there is a pressure for signal economy in language use—which is frequently identified as one of the prime causes or "motivations" for language change besides expressivity (e.g. by Zipf 1935; Langacker 1977; Keller 1994; Hopper and Traugott 2003; Croft 2000; Levinson 2000) and is often related to routinisation and frequency of use (e.g. Givón 1979; Bybee 1985, 2002; Haiman 1983, 1994)

The bias of language users towards signal economy can be explained by the fact that articulation constitutes a bottleneck for communication: it slows down the transmission of information. I have explained before that the transfer of information from one individual's mind to another's is only possible indirectly by means of modifying the shared physical environment (cf. section 3.1.1). I will refer to this process of modifying the physical environment as *articulation*—which is to be understood in the broadest sense here, including the production of speech signals just as well as the production of gestures or any other behaviour that serves as an ostensive stimulus. Levinson (1995:95f.) characterises articulation as a "relatively slow and inefficient process, which acts as a bottleneck in the entire communicative procedure." He claims that we can think faster than we can speak. As evidence, Levinson lists the observations that (i) we can perform complex planning tasks while we speak, (ii) we have no problems understanding pitch-corrected speech at double speed, and (iii) we are able to scan a printed page far faster than we could read it aloud. He concludes that "[p]sycholinguistic evidence [Levinson refers to Wheeldon and Levelt (1995) and Mehler et al. (1993)] seems to suggest that all the other processes in the entire complex chain of production and comprehension systems could run three or four times faster than the normal pace dictated by the articulation process" and that therefore "[t]he articulation bottleneck in human communication raises interesting questions from a design perspective: any trade-off from coded content to inferential meaning may greatly increase the speed of communication." (Levinson 1995:96).

However, the question this argumentation does not answer yet is why less time-consuming articulation should be advantageous to the goal of transferring information and therefore constitute an aspect of design for communication in the first place. If we want to say that a communication system that employs shorter signals is better designed for communication than an equally expressive one that uses longer signals because, on average, the former takes less time to transfer the same information than the latter, then we first have to explain why the speed at which some information can be transferred matters at all. What can be gained from being able to communicate faster? Two answers can be given to this question: a functional one and a non-functional one.

For one, the longer communication takes, the more likely it is to fail to achieve its goal. This may be due to three reasons. First, the environment may make it necessary that the respective information be available as quickly as possible. If we are hunting together, and I happen to see a gazelle that I want you to shoot, then I need to convey this information to you as long as the gazelle is still there. Or if I want to warn you of an approaching danger, I also need to make that danger manifest to you in time so that you can still escape it. If I communicate too slowly, I will not achieve my goal of making you shoot the gazelle or making you flee the impending danger. To use Austin's (1962) terms, my speech act will fail to have the intended perlocutionary effect. Second, by employing long and cumbersome signals, the communicator may be faced with the addressee's processing and memory limitations. Due to this, communication will, at least partly, fail. Third, if the transmission of information takes too much time, the communicator may loose the addressee's attention. As Sperber and Wilson (1995:vi) point out, "[t]o communicate is to claim an individual's attention." As a speaker, I have a very selfish interest in being fast because I want to convey some information to manipulate your behaviour, and I can only do so while I have your attention—which might fade quickly. In summary, all three mentioned points imply that the goal of a communicative act is more likely to be achieved the less time the communicative act consumes.

In addition, speakers (and hearers) may have an interest in shorter signals because they come at a lower cost. The fewer signals I need to produce, and the shorter they are, the less energy (and time) I will spend to achieve my goal. For the same reason, the addressee's attention span may be limited: paying attention comes with some cost, which the respective individual may attempt to minimise. The argument from cost is not functional with regard to communication: its motivation has nothing to do with the goal or perlocutionary effect that communication attempts to achieve—but minimising cost may itself be seen as a component of that goal (and in fact, as part of any goal envisaged by human behaviour).

The conclusion is this. A communication system whose average signal length is smaller is better adapted to the articulation bottleneck than one whose average signal length is higher. Because articulation is the slowest part of the communicative process, having shorter signals increases the speed at which speaker meanings can be transferred, which in turn makes it more likely that the respective communicative act is successful and at the same time minimises the cost at which the goal of the communicative act is achieved. A relatively low average signal length therefore adds to the appearance of design for communication.

I will now investigate the explanatory potential of the introduced model of language evolution with regard to this aspect of the appearance of design for communication. In particular, I will look at the role that pragmatic plasticity plays in the emergence of grammars that are better adapted to the articulation bottleneck. To this aim, I will first discuss how pragmatic plasticity can be conceived as offering "tools" to reduce signals and then study with the help of simulations how the average signal length of a grammar develops in the model and what role pragmatic plasticity plays in this process.

5.3.1 *Pragmatic plasticity as a "signal-reduction tool"*

I suggest that it is helpful to approach the role of pragmatic plasticity in the emergence of signal economy from a perspective informed by information theory: language users face a typical data-compression problem (Sayood 2006 offers a comprehensive exhibition of issues related to data compression and constitutes the standard text of this technical domain). Because of the articulation bottleneck, it is advantageous (for communication) to minimise the length of the employed signals *in individual speech acts*. The question is how a language user can *reduce* the length of the signal that his grammar makes available for an intended speaker meaning.

A first way to reduce the length of a signal is to drop some of its phonological material. That speakers do this has frequently been documented in studies of language change. Such behaviour has often been "fossilised" in a language: it then led to an erosion of the form and the permanent loss of the temporarily dropped phonological material, for example in cases of aphaeresis (the loss of a word-initial vowel as in *opossum* > *possum*), apocope (the loss of a word-final vowel or any word-final segment as in Old Spanish *pane* > modern Spanish *pan* 'bread') or syncope (the loss of an unstressed syllable between consonants as in Latin *septimana* > Spanish *semana* 'week').[11] A more complex example of such phonological loss, which we have already encountered in chapter 3, is the change of the English expression *going to* into *gonna*.

Phonological change is not part of the topic of this thesis and has consequently not been included in the introduced computational model. However, as a possible avenue for future research, I want to point out here that phonological change too can be interpreted as being initiated by pragmatic plasticity: not by an under- and/or overspecification of the *meaning* but by an under- and/or overspecification of the *form*. A speaker may underspecify the conventional linguistic form he is using by only producing as much phonological material as is necessary for the hearer to be able to identify it. Similarly, he may overspecify a conventional linguistic form by inserting ignorable phonological material where convenient. Such additional phonological material is ignorable as long as it does not turn the signal into some other conventional form. Such an analysis has the potential to identify a common cognitive base for both phonological and semantic change

[11] The examples stem from Trask (2000).

and complies with the argument that forms are meanings just as well, which I have presented in section 4.2.1.

This observation also points to the role that pragmatic plasticity, as the under- and/or overspecification of the speaker meaning, can play in the minimisation of the signal. Pragmatic plasticity offers "tools" to reduce signals *in specific contexts*. Underspecification allows a user to omit parts of the signal that can be inferred from the context. In example (4), for instance, the signal whose conventional meaning underspecifies the speaker meaning by leaving the second *old* out, (4a) is shorter than the signal whose conventional meaning fully specifies the speaker meaning (4b). Overspecification, on the other hand, allows for the use of shorter expressions with extra information that will be ignored in the given context. Such a case is illustrated in example (5): the metaphorical expression in (5a) is shorter than the literal description in (5b), even though it contains more semantic aspects than the latter.

(4) a. Most hearing aids are sold to old men and women.
    b. Most hearing aids are sold to old men and old women.

(5) a. Sally is a chameleon.
    b. Sally frequently changes her appearance.

In technical terms, the context-specific signal reduction provided by pragmatic plasticity constitutes so-called *lossy data compression*. Lossy compression techniques are used, for instance, to transmit audio and video signals (e.g. the MPEG compression format) or graphical information (e.g. the JPEG compression format). Data compression techniques are called lossy if the data that can be reconstructed from the signal they produce differ from the original data (Sayood 2006:5). This is the case for cases of under- and/or overspecification: the meaning recovered from the signal, the signal meaning, differs from the original message, the speaker meaning. Such a distortion is acceptable as long as it is not relevant to the environment or application for which the data compression has been carried out. The loss of some colour information in a picture, for instance, may not be relevant as long as the recipient can still recognise what the picture shows. Similarly, the distortion of speaker meaning through under- and/or overspecification is acceptable if the addressee is not confused or misled by the lack of information in underspecification or the extra information in overspecification, that is, as long as he is still able infer the intended speaker meaning on the basis of the encoded signal meaning. The important point about lossy data compression is that "[i]n return for accepting this distortion in the reconstruction, we can generally obtain much higher compression ratios than is possible with lossless compression" (Sayood 2006:5). Pragmatic plasticity thus equips language users with a set of particularly powerful methods of signal reduction.

In specific contexts, pragmatic plasticity offers tools to reduce signals; but note that this fact does not entail that the employment of pragmatic plasticity *always* results in signal minimisation. Expressions that under- and/or overspecify the speaker meaning are not in all cases shorter than their literal and fully specific counterparts. This fact can be illustrated with the following hypothetical examples, such as may occur in the introduced computational model. In the situation depicted in (6), underspecifying the intended speaker meaning *AB* by only producing the signal for *A* and having the addressee infer *B* from the context does not reduce but rather increases the length of the signal. In the situation shown in (7), overspecifying the intended speaker meaning *A* by producing the signal for *AB* does not reduce the signal either but also increases its length. Under- and overspecification do therefore not in all cases automatically lead to signal minimisation—their effect with respect to signal economy depends on the length of the available alternatives.

(6) *Underspecification without signal minimisation*

Input:	I-language:	$X \rightarrow AB$
		$YZ \rightarrow A$
	Speaker meaning:	*AB*
	Inferable meaning:	*B*
	Ignorable meaning:	–
Calculations:	Possible signal meanings:	*A*, *AB*
	Possible signals:	*YZ*, *X*

YZ (underspecification) $>$ *X (no underspecification)*

(7) *Overspecification without signal minimisation*

Input:	I-language:	$XY \rightarrow AB$
		$X \rightarrow A$
	Speaker meaning:	A
	Inferable meaning:	–
	Ignorable meaning:	B
Calculations:	Possible signal meanings:	A, AB
	Possible signals:	X, XY

XY (overspecification) > X (no overspecification)

What I am interested in now is the link between the fact that pragmatic plasticity provides tools for signal reduction in individual usage situations and the degree to which a whole grammar comes to be adapted to the articulation bottleneck. Through conventionalisation, individual usages enter the system and change it. Newly added form-meaning pairs may thereby reduce the average signal length of a grammar and thus increase the degree to which it exhibits the appearance of design for communication. I therefore formulate the general hypothesis that, because of pragmatic plasticity and conventionalisation, a communication system like language resembles an evolving lossy compression algorithm that constantly adapts to its environment. With this hypothesis, I explicate a less specific statement by Langacker (1977:106; cited also in Hopper and Traugott 2003:72), who says that "[i]t would not be entirely inappropriate to regard languages in their diachronic aspects as gigantic expression-compacting machines." It needs to be stressed, however, that the availability and use of pragmatic plasticity does not per se lead to a lower average signal length in a grammar since, as I have shown above, under- and overspecification do not always minimise the length of the produced signal. I will therefore employ computer simulations to investigate under what conditions the formulated hypothesis holds.

5.3.2 The development of signal economy in the model

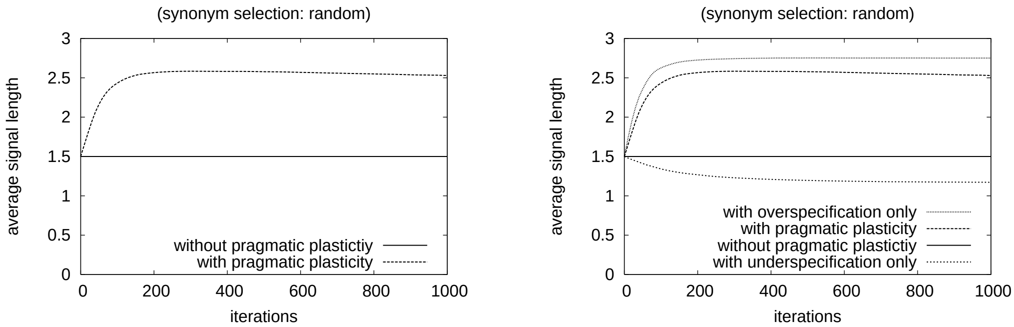
To study the role of pragmatic plasticity in how, and to what degree, symbolic communication systems come to adapt to the articulation bottleneck, I will compare the behaviour of the computational model with regard to signal economy in two series of simulations, one in which pragmatic plasticity is available and one in which the agent cannot employ pragmatic plasticity. Under the first condition, the agent is able to combine symbols and to use under- and/or overspecification to convey the intended signal meaning. Under the second condition, the agent will only be able to use extant symbols literally and to combine them but will not have access to under- and/or overspecification. First, I will look at the impact that the presence or absence of pragmatic plasticity has on the development of an already established grammar and then compare these observations with what happens when a grammar emerges anew from a non-symbolic starting point. A detailed description of how the average signal length of a grammar is measured in the simulations is provided in appendix B.2.

5.3.2.1 *Impact on already established grammars*

In order to investigate how the presence or absence of pragmatic plasticity can change the average signal length of an already established grammar, I employ simulations that start in a state where the agent already possesses a fully expressive grammar. Every potential speaker meaning can be expressed by the agent no matter whether pragmatic plasticity is available or not. The agent needs to be provided with a fully expressive grammar in these simulations so that the aspect of average signal length can be studied in isolation and the expressivity of the agent's grammar is prevented from influencing the result.

I start my investigation with a simple, idealised setup similar to the one used previously to study expressivity. The details of this setup are given in (8). The agent's semantic space is made up of the four semantic units *A*, *B*, *C* and *D* (which all occur with equal probability), meanings maximally consist of two semantic units, contexts contain zero or one inferable or ignorable semantic unit, there is no loss (i.e. the entrenchment threshold is set to 0). I will vary the strategy the agent employs to select synonyms because this factor has a direct impact on the length of the signals in individual usage situations and can therefore also be hypothesised to influence the way in which the average signal length of the whole grammar develops over time.

(8)	Semantic units:	*A*, *B*, *C*, *D*
	Maximal meaning complexity:	2 semantic units
	Maximal context size:	1 semantic unit
	Entrenchment threshold:	$t = 0.0$



(a) (b)

Figure 5.15: Development of the average signal length of an already established grammar (synonym selection: random). Each line represents the average over 100 independent simulation runs under the conditions listed in (8) and with the initial grammar given in (9). If signals are selected from synonyms at random, pragmatic plasticity (with both under- and overspecification being available) is not advantageous with regard to the development of the signal economy of the agent's grammar.

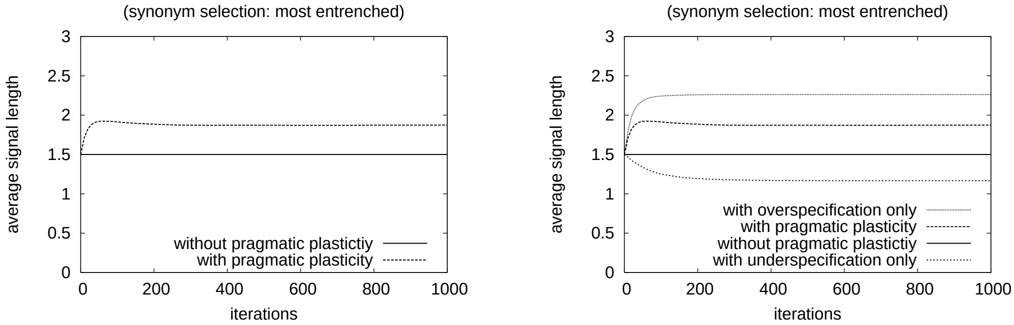
In (9), I provide the grammar the agent is equipped with at the beginning of the simulation runs. This grammar is fully expressive, that is, it can express all combinations of semantic units that are possible.

(9) Initial grammar:
$W \rightarrow A$
$X \rightarrow B$
$Y \rightarrow C$
$Z \rightarrow D$
$WX \rightarrow AB$
$WY \rightarrow AC$
$WZ \rightarrow AD$
$XY \rightarrow BC$
$XZ \rightarrow BD$
$ZY \rightarrow CD$

Synonym selection: random. Fig. 5.15(a) compares the development of the average signal length of the agent's grammar with and without the availability of pragmatic plasticity—the latter serving as the NULL-hypothesis—under the condition that the agent always selects randomly if more than one signal is possible in a situation. What Fig. 5.15(a) shows is that the occasional use of pragmatic plasticity leads to a higher average signal length than the continuous literal use of extant conventions. If the agent does not use pragmatic plasticity, the average signal length remains stable at its original level. But if the agent uses pragmatic plasticity, the average signal length of the agent's grammar is increased and stabilises at a higher level. Under these conditions, the availability of pragmatic plasticity is thus not advantageous with regard to signal economy—which runs contrary to our hypothesis.

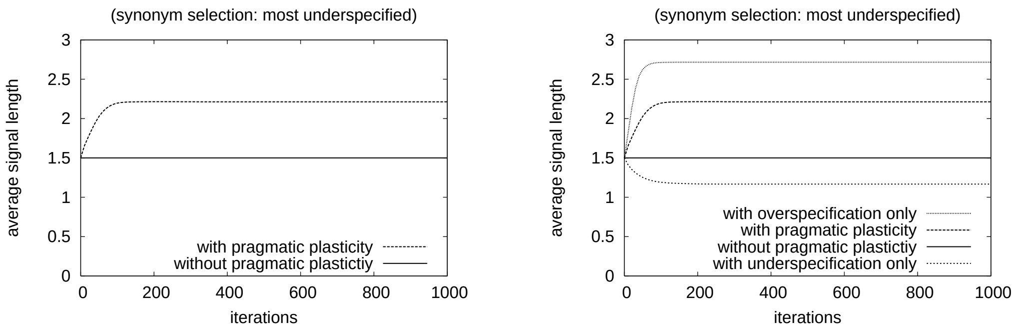
The reason for this result becomes more evident if the impact of underspecification and overspecification are analysed separately. Fig. 5.15(b) represents the same as 5.15(a) but additionally plots the development of the average signal length of the agent's grammar under the conditions that (i) the agent can make use of underspecification but not of overspecification, and (ii) that he only has overspecification but not underspecification available. It can be observed that if only underspecification is used occasionally but not overspecification, the average signal length of the agent's grammar decreases over time and stabilises at a level that is lower than the one at which it remains if no pragmatic plasticity is employed at all. In contrast, if the agent only uses overspecification but never underspecification, the average signal length of the grammar stabilises at a level that is even higher than if he makes use of both. It can thus be said that, under the given conditions, underspecification tends to reduce the average signal length while overspecification increases it. In the example, the occasional use of overspecification is enough to undo the positive effect that underspecification has on the average signal length of the grammar.

Synonym selection: most entrenched. The same phenomenon can be observed in simulation runs where the agent consistently chooses the most entrenched available signal. Figs. 5.16(a) and 5.16(b) plot the results for this case. The overall behaviour of the model is the same as for the cases where synonyms are chosen at random. The availability of pragmatic plasticity is not advantageous with respect to the average signal length of the agent's grammar because the positive effect of underspecification is outweighed by the negative impact of the occurring cases of overspecification. The only difference between the results produced by the two selection strategies is that if the most entrenched synonyms are chosen, the impact of overspecification is somewhat less than if synonyms are selected at random. While the use of pragmatic plasticity still leads to an increase of the average signal length of the grammar, that increase is not as big as in simulation runs with random synonym selection.



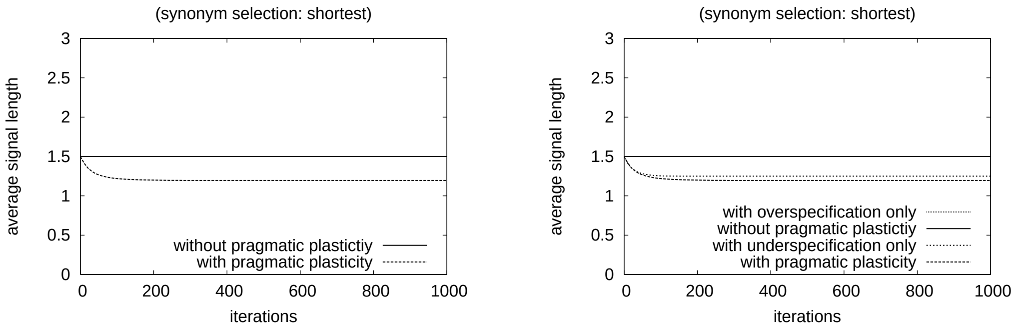
(a) (b)

Figure 5.16: Development of the average signal length of an already established grammar (synonym selection: most entrenched). Each line represents the average over 100 independent simulation runs under the conditions listed in (8) and with the initial grammar given in (9). If the most entrenched synonym is chosen, pragmatic plasticity (with both under- and overspecification being available) is not advantageous with regard to the development of the signal economy of the agent's grammar.



(a) (b)

Figure 5.17: Development of the average signal length of an already established grammar (synonym selection: most underspecified). Each line represents the average over 100 independent simulation runs under the conditions listed in (8) and with the initial grammar given in (9). If the most underspecified synonym is chosen, pragmatic plasticity (with both under- and overspecification being available) is not advantageous with regard to the development of the signal economy of the agent's grammar.



(a) (b)

Figure 5.18: Development of the average signal length of an already established grammar (synonym selection: shortest). Each line represents the average over 100 independent simulation runs under the conditions listed in (8) and with the initial grammar given in (9). If the shortest synonym is chosen, pragmatic plasticity (with both under- and overspecification being available) *is* advantageous with regard to the development of the signal economy of the agent's grammar.

Synonym selection: most underspecified. Even if the agent consistently chooses the most underspecified signal, as some pragmatic theories might suggest (cf. section 4.1.2.3), pragmatic plasticity is still not advantageous with regard to the degree in which the grammar comes to be adapted to the articulation bottleneck. Fig. 5.17 plots the development of the average signal length under this condition. Although the agent constantly chooses the most underspecified signal, the negative effect of the occurring cases of overspecification, which apparently result in longer signals, cannot be undone. Its impact even seems to be somewhat stronger than in the earlier simulations. The strategy to choose the signal that underspecifies the speaker meaning most does therefore not reliably lead to a decreased average signal length.

Synonym selection: shortest. The picture is different if the agent consequently chooses the shortest available signal—which is what the presence of the articulation bottleneck would suggest in the first place. Fig. 5.18 shows that under this condition, pragmatic plasticity is advantageous and that its occasional employment leads to a decrease of the average signal length of the agent's grammar. Pragmatic plasticity is now only used if it really results in a reduction of the signal length in the given usage situation: it thus unfolds its positive potential without being hindered by the negative impact of cases of under- or overspecification that produce signals that are longer than the ones that already existed.[12]

In summary, the simulations discussed suggest that the strategy employed to choose a signal from a range of available synonyms can have a substantial impact on whether the availability of pragmatic plasticity is advantageous with regard to the overall signal economy of a symbolic communication system. Under the simulated simple and idealised conditions, the occasional use of pragmatic plasticity reduces the average signal length of an already established grammar—and thus increases the degree to which that grammar is adapted to the articulation bottleneck—only if the agent chooses the shortest signal whenever he faces a situation in which more than one option is possible. In all other cases (i.e. where the agent chooses at random or the most entrenched signal or the signal that underspecifies the speaker meaning the most), the occasional positive effect of underspecification is outweighed by the occasional negative effect of overspecification. It remains to be tested if the picture is the same if the agent starts from a non-symbolic state, that is, if he is not given an already established grammar at the beginning of the simulation run.

5.3.2.2 Impact on emerging grammars

I will now address the question of whether the nature of the initial state has an impact on the development of signal economy. Specifically, I will ask how pragmatic plasticity influences the average signal length of a communication system that emerges from "nothing," that is, from a state where no conventional form-meaning association has been established yet. Because I want to be able to compare the results to those found for the development of signal economy in an already established grammar, I will use the same basic parameter set as before.

Even though the agent is now not equipped with an established grammar at the beginning of the simulation runs, we have to make sure that full expressivity can be reached both if pragmatic plasticity is available and if it is not available. This means that the agent needs to be able to express all possible semantic units by means of producing direct evidence. Otherwise, the agent could not cover

[12]Note that in the given example, overspecification appears to lead to longer signals in practically all cases in which it occurs. If the agent constantly chooses the shortest signal, simulations in which overspecification but not underspecification is available lead to the same result as those in which pragmatic plasticity is not available at all. Under the simulated conditions, literal use therefore always produces shorter signals than an alternative which overspecifies the speaker meaning.

the whole semantic space in the condition without pragmatic plasticity. At the beginning of a simulation run, the agent thus possesses the set of producible cues listed in (10). With these cues, the agent can communicate every possible meaning either by producing a single cue or by combining two or more of the cues—pragmatic plasticity does not need to be resorted to.

(10) Initial "grammar": $A \rightarrow A$
$B \rightarrow B$
$C \rightarrow C$
$D \rightarrow D$

We can now compare how pragmatic plasticity influences the signal economy of a newly emerging grammar with how it influenced the signal economy of an already established grammar. Fig. 5.19(a) shows again the development of the average signal length of an established grammar; Fig. 5.19(b) gives the respective development for a newly emerging grammar. In both sets of simulations, the agent chose synonyms at random. We can observe that in both cases, signal economy stabilises at about the same level. The only noticeable difference is that where the agent has to start without a grammar, signal economy has to start from the bottom.

We get roughly the same picture if the agent chooses the most underspecified signal, as shown in Figs. 5.19(c) and 5.19(d) respectively, and if he chooses the shortest signal, as shown in Figs. 5.19(e) and 5.19(f). In all these cases, the average signal length of the newly emerging grammar stabilises more or less at the same level as it does for an already established grammar. More precisely, we can distinguish between *quantitative* differences and *qualitative* differences. We can speak of a quantitative difference if the average signal length is higher in one case than in another and of qualitative difference if the availability of pragmatic plasticity is an advantage with regard to signal economy in one case but a disadvantage in the other. While there are minor quantitative differences, there is no qualitative difference in the three pairs of figures: both for established as well as for newly emerging grammars, the availability of pragmatic plasticity is advantageous if the agent consistently chooses the shortest signal but not if he chooses at random or always selects the signal that underspecifies the speaker meaning most.

The situation is different, however, if the agent selects the most entrenched signal. Under this condition, we do get a qualitative difference: Fig. 5.20 shows that

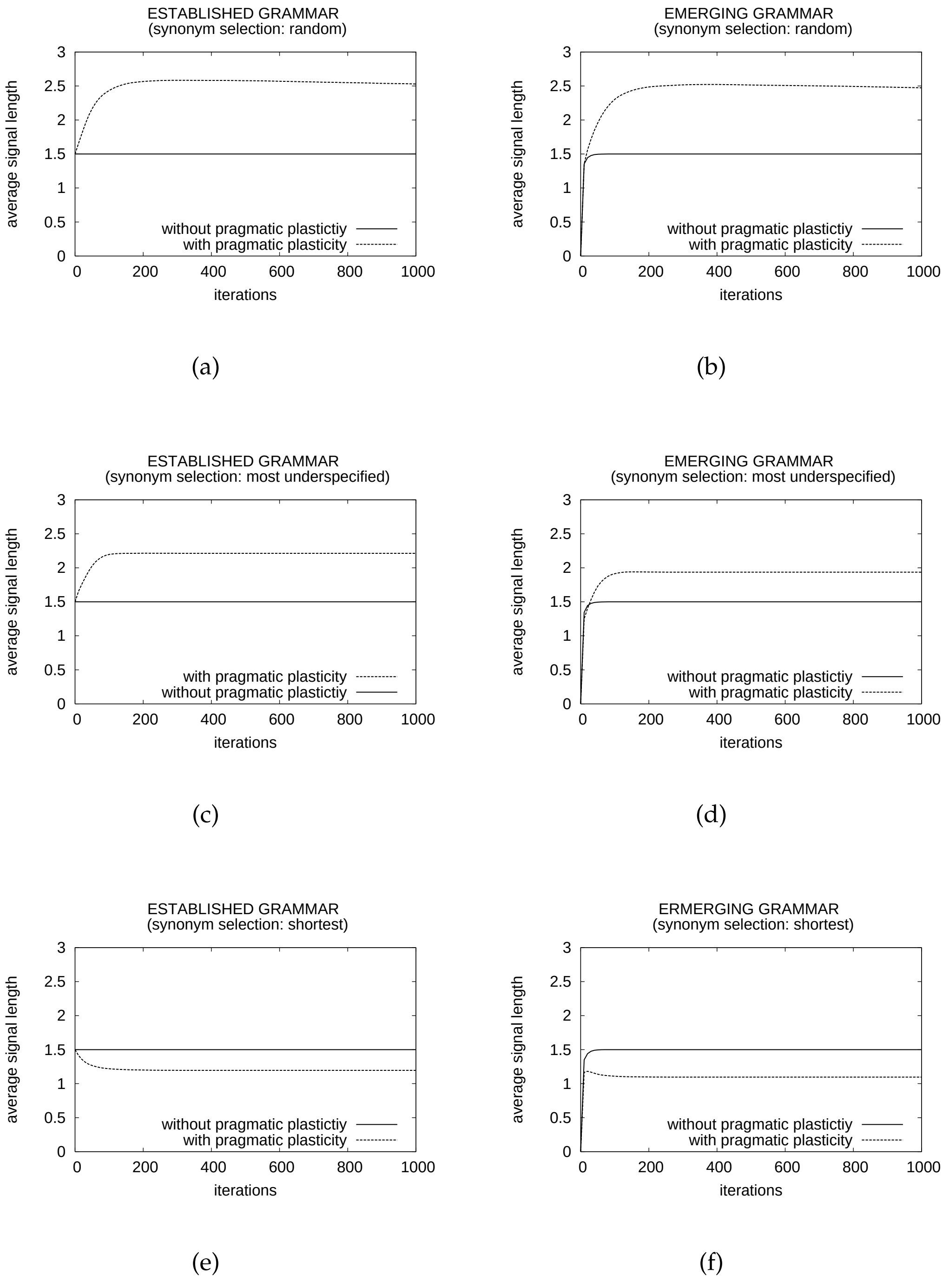
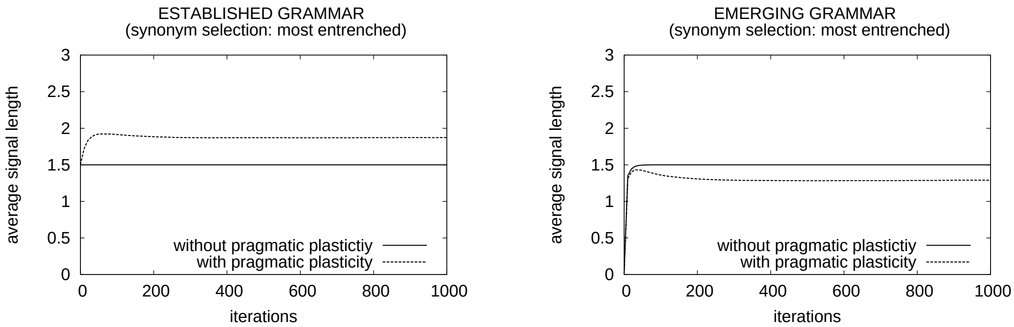


Figure 5.19: Development of the average signal length of a newly emerging grammar with synonym selection strategies: (a–b) random; (c–d) most underspecified; (e–f) shortest. The figures in the left-hand column plot the development of an already established grammar; the figures in the right-hand column the development of a newly emerging grammar. Each line represents the average over 100 independent simulation runs under the conditions listed in (8) and with the initial grammars given in (9) and (10) respectively. Both for established as well as for newly emerging grammars, pragmatic plasticity is advantageous with selection strategy "shortest" and disadvantageous with selection strategies "random" or "most underspecified."



(a) (b)

Figure 5.20: Development of the average signal length of a newly emerging grammar (synonym selection: most entrenched): (a) development of an already established grammar; (b) development of a newly emerging grammar. Each line represents the average over 100 independent simulation runs under the conditions listed in (8) and with the initial grammars given in (9) and (10) respectively. Pragmatic plasticity is disadvantageous with regard to signal length if the simulation starts with an already established grammar but is advantageous if it begins in a state where the agent does not possess conventionalised form-meaning pairings yet.

the occasional use of pragmatic plasticity is a disadvantage for signal economy in an established grammar but it leads to a decreased average signal length in a newly emerging grammar. It thus depends on the initial grammar whether pragmatic plasticity proves to be advantageous to the development of signal economy or not. What is relevant for our purposes is that in the typical evolutionary situation, that is, where users start without any form of symbolic communication system, the strategy of always choosing the most entrenched signal proves to yield a better adaptation to the articulation bottleneck if pragmatic plasticity is available than if it is not. Note that under the described circumstances (i.e. a non-symbolic initial state and the availability of pragmatic plasticity), the average signal length of a grammar remains low even though the user does not strive to minimise signals in individual usage events but simply exhibits whatever behaviour is triggered first.[13]

[13]The impact of loss, meaning complexity and the probability structure of the semantic space on signal economy were investigated too. However, under the described conditions, these factors only yielded quantitative but no qualitative differences.

5.3.2.3 *Summary and interpretation*

With the help of the simulations discussed, I have shed some light on the role that pragmatic plasticity plays in how, and to what degree, a symbolic communication system adapts to the articulation bottleneck. Language does not exhibit pragmatic plasticity because of the pressure for signal economy in language use, it rather exhibits pragmatic plasticity because linguistic communication is ostensive-inferential. However, the simulations discussed here have illustrated that the fact that pragmatic plasticity is available can, under the right conditions, help a language adapt to the articulation bottleneck by providing it with tools for context-specific signal minimisation. Specifically, the use of pragmatic plasticity reduces the average signal length of a communication system (i) if users are biased towards choosing the shortest signal, and, in case the system starts from zero, also (ii) if they are biased towards choosing the most entrenched signal. It is particularly relevant that pragmatic plasticity proves to be advantageous with regard to signal economy specifically with these two selection strategies since these two strategies are most plausible from a psychological perspective and accommodate the bias for an economy of effort that can be assumed for any behaviour.

5.4 Ambiguity

In the last part of this chapter, I will turn to a putatively dysfunctional property of language: ambiguity. Ambiguity describes any state where a linguistic code contains forms that are conventionally associated with more than one meaning. The perception that ambiguity is a dysfunctional feature, and that its ubiquity in language therefore constitutes an evolutionary puzzle, is based on the idea that an optimal code must be transparent. According to Langacker (1977:110), transparency denotes a type of linguistic optimality that is "based on the notion that the ideal or optimal linguistic code, other things being equal, will be one in which every surface unit (typically a morpheme) will have associated with it a clear, salient, and reasonably consistent meaning or function, and every semantic element in a sentence will be associated with a distinct and recognizable surface form." Langacker (1977:110) further specifies that "[l]anguages are thus optimal along this parameter to the extent that they show a one-to-one correspondence between units of expression and units of form" and concedes that "[o]bviously, though, natural languages routinely and massively fail to achieve this state." These statements exemplify the wide-spread notion that ambiguity constitutes a dysfunctional feature in the sense that a non-ambiguous code would be better suited for communication. I will question this assumption.

It is often overlooked that the putative presence of a dysfunctional feature allows for more than one possible conclusion. First, it may be an indicator that the respective entity does not exhibit design for the assumed function, that is, that the wrong function has been assumed. According to this reasoning, the fact that language is ambiguous, and that ambiguity is dysfunctional with regard to communication, implies that language is not the product of an adaptation to its use for communication. This is the line of argumentation advocated by Chomsky. In defence of his non-selectionist account of language evolution, he argues explicitly that ambiguity constitutes a serious problem to the assumption that language exhibits design for communication:

> The natural approach has always been: is it [i.e. language] well designed for use, understood typically as use for communication? I think that's the wrong question. The use for communication might turn out to be a kind of epiphenomenon. I mean, the system developed however it did, we really don't know. And then we can ask: how do people use it? It might turn out that it is not optimal for some of the ways in which we want to use it. If you want to make sure that we never misunderstand one another, for that purpose language is not well designed, because you have such properties as ambiguity. (Chomsky 2002:107)

Chomsky's conclusion is consequently that language is not well designed for its communicative function but that it *is* well designed "with regard to the internal system with which it must interact" (2002:108). However, he fails to specify in what way ambiguity fits into this picture. But, more fundamentally, the argument that the presence of a dysfunctional feature indicates that the entity in question does not exhibit design for the assumed function is fallacious for a basic reason: it is based on the wrong assumption that adaptations are perfect. Design does not equal perfection (e.g. Dawkins 1999:ch. 3, Kinsella in press:ch. 2). An engineer, for instance, may design a tool to perform a specific task but he may not have found the best possible solution to the problem posed by the task yet. Nevertheless, the tool will exhibit design for the respective function, at least a certain degree of it. The same holds for biological systems. Futuyma (1998:356–361), for instance, emphasises that, while design is one means of recognising an adaptation brought about by natural selection, one thing not to expect of natural selection and adaptation is perfection. We can presume that this is also true for entities that are the product of cultural evolution. Therefore, the presence of ambiguity in language does not mean that language does not exhibit design for communication.

A second way of explaining the presence of an apparently dysfunctional feature is to assume that the entity adapts to more than one function, and that the feature in question is a product of this second adaptive process. Some studies carried out within the framework of the iterated learning model, for instance, argue that certain linguistic structures can be explained by assuming that language appears to be designed to be learnable (e.g. Kirby 2002a; Kirby and Hurford 2002; Smith 2003c; Brighton et al. 2005). Within this line of research, I have shown that syntactic ambiguity can emerge from iterated grammar induction and can thus be interpreted as a side-effect of a grammar's adaptation to the learning bottleneck (Hoefler 2006b).

In this chapter, however, I want to pursue a third possibility, namely the one that the feature in question may only *appear* to be dysfunctional and that rather the assumptions about *how* the respective function is fulfilled are faulty or incomplete. I argue that ambiguity only constitutes a dysfunctional property of language if one works with an over-simplified model of communication: the so-called code model (cf. section 3.1.2). Ambiguity is only detrimental to successful communication if disambiguation in context is not possible. But as soon as one recognises that the process of encoding and decoding is enhanced with inference from context—or even assumes an ostensive-inferential model of communication in the first place, as I have done in chapter 3—ambiguity does not pose an insurmountable problem to communication anymore. Coding then only represents one component of the production and comprehension process, and even though at some stage the produced form may have the potential to lead to more than one *signal meaning*, only one *speaker meaning* will result in the end (that is, if the provided clue is sufficient). In light of the fact that in order to establish communication, two individuals have to recognise common-ground knowledge of what information is relevant in the given context, ambiguity thus merely appears as *non-functional* rather than *dysfunctional*: it neither facilitates nor seriously hinders communication. This analysis is supported by the fact that ambiguity is ubiquitous in language and humans do not even seem to notice its presence for most of the time. Traugott and Dasher (2005:12), for instance, point out that there is no evidence that speakers actually strive for a minimisation of ambiguity in the linguistic code:

> While the demands of an increasingly technological society favor the development of rigidly specified lexical distinctions, for example, among diseases, legal rulings, or linguistic terminologies, very few, if any, words in ordinary language have only one interpretation.

The perception of ambiguity as a dysfunctional feature seems to be based on a failure to distinguish between signal meaning and speaker meaning. An ambiguous code leads at times to more than one signal meaning. However, this fact does not mean that communication needs to fail as the context in which the respective usage event happens may still allow for only one speaker meaning to be inferred. Ambiguity in the signal meaning does not necessarily lead to ambiguity in the speaker meaning. Pinker and Bloom (1990:713, emphasis added) consequently note that language exhibits design for communication because it allows for "minimising ambiguity *in context*" (as opposed to a minimisation of ambiguity in the code). In summary, we can therefore state that if one assumes an ostensive-inferential model of communication, ambiguity in the code (within limits; I will return to the problem of massive ambiguity below) appears merely as a non-functional rather than a dysfunctional feature.

The analysis just presented takes a synchronic perspective on the question of design for communication. I will now look at the problem from a diachronic perspective, and go a step further and argue that, from a diachronic perspective, ambiguity even appears as a positively *functional* feature.

5.4.1 Ambiguity as a functional feature

While ambiguity may not directly contribute to the successful establishment of communication in individual usage situations, I will show that ambiguity plays a crucial role in the adaptability of a code to its communicative function in the course of its cultural evolution. In order to make a case for ambiguity as a functional feature, I will argue that a code that allows for ambiguity has a higher adaptability to the functional pressures for (i) maximised expressivity and (ii) minimised average signal length. Ambiguity thus directly contributes to the two aspects of the appearance of design for communication that I have discussed in the previous parts of this chapter.

5.4.1.1 *Its relevance for expressivity*

To understand the role that ambiguity plays with regard to the expressivity of a code, we need to re-examine how meanings change in the course of iterated language use. As I have detailed in chapter 3, extant signals are used in novel (i.e. under- and/or overspecified) ways, and these novel usages become associated with the employed signal: the signal acquires a new meaning. This meaning was potentially not expressible by the user's grammar before. Traugott and Dasher (2005:11) accurately summarise the role that ambiguity comes to play in this process:

> Semantic change cannot be studied without drawing on a theory of polysemy because of the nature of change. Every change, at any level in a grammar, involves not "$A > B$," i.e. the simple replacement of one item by another, but rather "$A > A \sim B$" and then sometimes "$> B$" alone.

Thus, when an extant symbol is used in a way in which it either under- and/or overspecifies the actually communicated meaning, and, as a consequence of this, a new association between the form of that symbol and the meaning it conveyed is established, the association between that form and its original meaning does not automatically disappear.[14] What emerges is rather a situation that historical linguistics refer to as *layering*: the co-existence of two variants of a symbol, an older and a newer one. Layering can designate a situation where two symbols have the same meaning but their form varies, or it can, as in our case, denote a situation of polysemy, where the form is the same but the meaning varies. The older meaning may eventually fall out of use, and this is usually the moment where the result of the process is perceived as semantic change rather than simple variation. However, as Traugott and Dasher (2005:11f.) point out "despite

[14]Traugott and Dasher's schema "$A > A \sim B\,(> B)$" does not specify individual types of semantic change. In section 3.2.2, I have described that pragmatic plasticity can bring about (i) semantic narrowing or specialisation through underspecification, (ii) semantic broadening or generalisation through overspecification, or (iii) a combination of the two if an extant symbol is used in a way that both under- and overspecifies the speaker meaning. These three general cases can be represented as instantiations of Traugott and Dasher's schema if the structure of the involved meanings is made explicit. In the following sub-schemata, which instantiate the general schema "$A > A \sim B\,(> B)$," meanings A and B are represented as combinations of one or more conceptual units C: (i) underspecification and narrowing: "$C_1 > C_1 \sim C_1C_2\,(> C_1C_2)$"; (ii) overspecification and broadening: "$C_1C_2 > C_1C_2 \sim C_2\,(> C_2)$"; (iii) under- and overspecification combined: "$C_1C_2 > C_1C_2 \sim C_1C_3\,(> C_1C_3)$." A fourth, limiting case is represented by situations of maximal under- and overspecification (cf. section 3.1.3): "$C_1 > C_1 \sim C_2\,(> C_2)$."

what is often thought, the loss of an earlier meaning is relatively rare. What is typical is the accretion of more and more meanings over time, hence the parenthetical status of the lone survivor B in the schema $A > A \sim B\,(> B)$." It is, of course, exactly this "accretion" of meanings resulting from polysemy that leads to the cumulative expansion of expressivity—and thus ultimately to better design for communication. Moreover, if there was no polysemy, that is, if newer meaning did replace older meanings, there would be semantic change but in effect no *increase* in expressivity: every time a new meaning would become expressible, an old meaning could not be expressed anymore. In conclusion, it can therefore be stated that ambiguity is not only not dysfunctional but it actually contributes to the emergence of better design for communication.

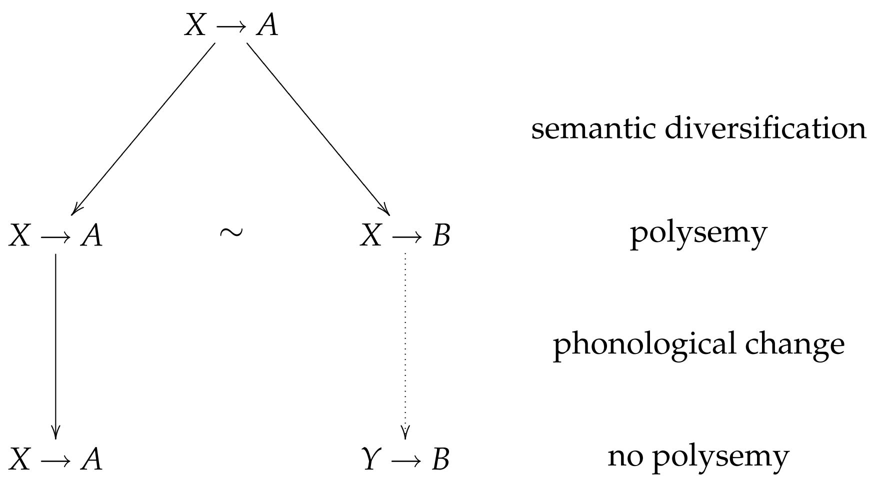
Without any differentiation in the forms of polysemous symbols, the scenario that has been introduced so far could be taken ad absurdum: it would, in its most extreme case, allow for one form to become associated with all possible meanings—or at least with all meanings that can be reached from it by means of pragmatic plasticity over time. However, the more meanings such a form would come to express, the less likely it would actually be that disambiguation would still be possible in context: the code would reach a *degree* of ambiguity that would indeed be dysfunctional. This problem is prevented by two processes. First, not only one pre-linguistic cue can become conventionalised. It is very likely that, even though there is one hypothetical "first symbol," other pre-symbolic cues would also become associated with the meanings they conveyed, so that very quickly, there would be a whole number of symbols—each with a different form—that could be used as stepping stones to explore new meaning spaces.

However, even this fact does not account for the extreme diversity of form that can be found in human language. A second process must therefore be taken into account: phonological change. In section 3.3, I argued that we need to assume the causal chain of change represented in (11) for phenomena of grammaticalisation.

(11) semantic change ⇒ change in frequency (and environment) of use ⇒ phonological change

This process can, in general, bring about a diversification not only of the meanings that a symbolic communication system can express but also of the forms it employs to do so. It can be applied to any form of semantic diversification, not just to those linguistic cases usually classified as grammaticalisation phenomena. I thus extend Traugott and Dasher's abstract schema as shown in (12).

(12)



A form X is originally associated with a meaning A. Through under- and/or overspecified use and subsequent conventionalisation, a situation of polysemy emerges: X is now associated with the original meaning A as well as with a new meaning B. Because B differs in frequency of occurrence from A, and because it is also used in different linguistic environments, the form that signifies B undergoes phonological change but the form that denotes A does not. The effect of this process is that the earlier instance of polysemy disappears.[15]

It needs to be noted, however, that phonological change happens according to the same schema that Traugott and Dasher (2005:11) apply to semantic change: phonological change too involves not "$X > Y$," that is, the replacement of one form with another, but rather "$X > X \sim Y\ (> Y)$." McMahon (1994:50), for instance, provides the following scenario for sound change:

> If we assume for the moment an abstract change of some phoneme X to some other phoneme Y, certain morphemes will undergo the change directly, but in others, pronunciation will fluctuate for a time, for individuals and/or for the community. This period of variation is characterised by the existence of doublets, morphemes with two pronunciations.

We can thus entertain the possibility of layering for phonological change too: that the old and the new form co-exist beside each other until one falls out of use. This

[15]Note, however, that phonological change can also introduce new instances of ambiguity, namely if the new form happens to coincide with an already existing form. Ambiguity that emerges through such coincidences is referred to as "homonymy." An example are the two English words that have the form *pen*, one denoting an enclosure and originating from the respective Old English word *penn*, the other referring to a writing utensil and being derived from an Old French word for 'feather', *penne*.

is the reason why in schema (12), the line connecting $X \rightarrow B$ and $Y \rightarrow B$ is only dotted: it indicates that only the result of this two-step process is shown, namely X in effect having turned into Y, but not the intermediate step of layering where X and Y co-exist alongside each other.

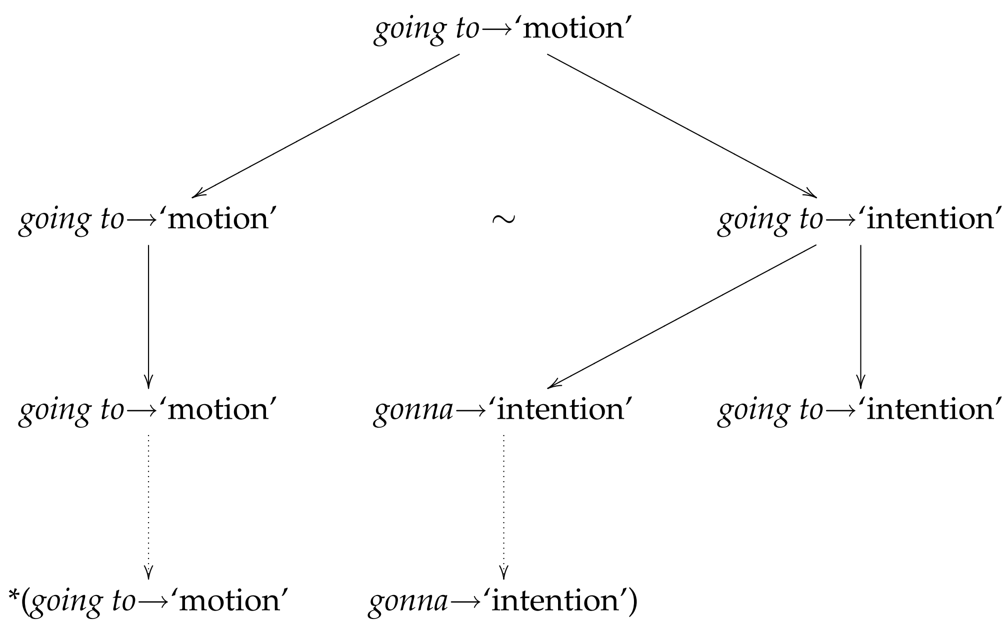
To illustrate how layering can occur as a result of semantic as well as phonological diversification, I refer once more to the example of the grammaticalisation of *be going to*. The state that this expression currently has in many English dialects exhibits semantic layering (*be going to* can be used to express motion as well as intention) as well as phonological layering (one of its meanings, 'intention', can be expressed by the older form *be going to* as well as the newer variant *gonna*). The already introduced examples (13) and (14) document this point.

(13) a. Are you *going to* church now?
     b. ∗ Are you *gonna* church now?

(14) a. No, I am *going to* stay here.
     b. No, I am *gonna* stay here.

Schema (15) visualises the development of *be going to*.

(15)



An initial semantic diversification leads to a state of polysemy ('motion' vs. 'intention'; refer to section 3.3.2 for a detailed discussion of this step). Due to its altered frequency and environment of use, the new form-meaning pair subsequently develops an alternative form (*gonna*). The last part of the schema, plotted with dotted lines, represents a hypothetical future stage at which polysemy has disappeared because 'intention' has come to be expressed exclusively by the new form *gonna*.

As I have already pointed out in section 5.3.1 above, phonological change is not dealt with in any further detail in this thesis. For the present purposes, it will suffice to state that it plays a vital role in a model of how symbolic communication systems grow more expressive because it provides a solution to the described problem of massive ambiguity. It is, however, important to emphasise that phonological change is not goal-directed in the sense that users would change forms to be able to distinguish between the individual meanings of an otherwise polysemous form. Under the circumstances considered here, forms change because their frequency of use and the environment in which they are used have changed as a result of a preceding semantic change. Phonological change of this type is linked to processes of automatisation (Givón 1979; Haiman 1994; Bybee 1998; Bybee and Thompson 2000).

The bottom line of the discussion that I have presented here is this: ambiguity is a necessary prerequisite for the expressivity of a symbolic communication system to increase in the course of cumulative cultural evolution through iterated under- and/or overspecified use. Without the possibility for ambiguity in the form of layering, pragmatic plasticity could not unfold its adaptive potential: extant conventions could not be used as stepping stones to incorporate neighbouring meaning spaces into the linguistic code and language would not become as well designed to meet the communicative needs of its users as it is.

5.4.1.2 *Its relevance for signal economy*

The advantage that ambiguity bears with regard to signal economy can be explained by a specific information-theoretic observation: the minimal average codeword length of a code that allows for ambiguity is smaller than the minimal average codeword length of a code that needs to express the same number of meanings with the same number of phonological units available but does not allow for ambiguity. This information-theoretic fact can be illustrated with the example codes shown in Table 5.3. Each of these codes needs to express three meanings (*a*, *b* and *c*) by means of two units of form (0 and 1) under the condition that only a specific degree of ambiguity is allowed. Underneath each code,

	Code 1	Code 2	Code 3	Code 4
a	0	0	0	0
b	10	10	1	1
c	11	01	01	1
ACL$_{min}$	1.67	1.67	1.33	1.00

Table 5.3: Ambiguous codes have a lower average codeword length. The table shows four types of codes of minimal average codeword length that express the meanings *a, b, c* with only two units of forms (0 and 1). Code 1: prefix code; Code 2: non-ambiguous non-prefix code; Code 3: ambiguous non-prefix code (individual codewords distinct); Code 4: ambiguous non-prefix code (individual codewords not distinct). The minimal average codeword length (ACL_{min}) decreases the more clearly a code allows for ambiguity.

the table gives the minimal average codeword length that can be achieved under the respective condition. Code 1 is a non-ambiguous code of minimal average codeword length. Codes of this type are also referred to as "prefix codes." For a code to be a prefix code, no codeword must occur as the first part of another, longer codeword. Code 2 is a non-prefix code: its codeword for meaning *a* (0) is also the initial part of the codeword for *c* (01). This code is still non-ambiguous but a parser may face a temporary situation of ambiguity: after processing 0 as the first unit of form of a longer sequence, one cannot know yet whether this denotes *a* or the first half of the codeword for *c*. Sequences that contain such temporal ambiguity occur in language in the form of so-called garden-path sentences like (16).

(16) The horse chased past the barn fell.

Code 3 exhibits ambiguity. No two codewords are the same but the codeword for *c* (01) cannot be distinguished from a sequence encoding *ab*. Structural ambiguity like this is ubiquitous in language—one example for it is attachment ambiguity as in (17).

(17) The policeman saw the spy with the telescope.
[The policeman saw the spy, who had a telescope.]
[The policeman, who used a telescope, saw the spy.]

The advantage of this condition, as opposed to the non-ambiguous variants in codes 1 and 2, is that its minimal average codeword length is lower. Code 4, finally, exhibits ambiguity to the degree that even individual codewords are the same: the codeword for *b* (1) is the same as the codeword for *c*. Under this condition, the minimally possible average codeword length is again lower and can be reduced to 1.

In summary, the presented examples illustrate the following fact: if the number of meanings that a code needs to express (here *a, b, c*) is higher than the number of the available units of form (here 0 and 1), then the average codeword length that is minimally necessary can be lowered by allowing for ambiguity. Given the high number of meanings humans want to express and the comparatively low number of units of form they can use, their signals would be considerably longer if language did not allow for ambiguity. Ambiguity facilitates greater signal economy and therefore constitutes a feature that is advantageous to the communicative function of language.

We can conclude that ambiguity as a *feature* is functional: its availability is advantageous to the degree to which the expressivity and the signal economy of a code can adapt to its environment in the course of its cultural evolution. It is only the *degree* of ambiguity that can potentially be dysfunctional if it is too high. The introduced computational model does not provide a way of determining at what degree ambiguity becomes dysfunctional. However, what we *can* study with the help of the model is if there are circumstances under which the emerging degrees of ambiguity are generally lower and thus less likely to be dysfunctional.

5.4.2 The development of ambiguity in the model

The lower the degree of ambiguity in a grammar is the less likely we are to face the problem of massive ambiguity that can be detrimental to successful communication. I have already pointed out in section 5.4.1.1 that the degree of ambiguity may be reduced by phonological change. With the following computer simulations, I study what other factors can lower the degree of ambiguity in an emerging and evolving symbolic communication system. I will first discuss the general behaviour of the model in idealised conditions. On the basis of the observations made, I will then formulate a number of hypotheses about the factors influencing the development of ambiguity in the course of cultural evolution, and I will test these hypotheses in subsequent simulation runs.

In appendix B.3, I provide a detailed description of how the ambiguity of the agent's grammar is measured in the simulations. Here, it will suffice to state that the applied definition of ambiguity takes into account the following aspects: (i) the proportion of constructions in a grammar that are ambiguous, and, for each ambiguous construction, (ii) the number of meanings that the same form is associated with (a form with three meanings is more ambiguous than a form with only two meanings) and (iii) the relative entrenchment of the individual meanings of an ambiguous form (a form is more ambiguous if both its meanings are equally entrenched and less ambiguous if one of its meanings is very salient while the other one is hardly entrenched at all).

5.4.2.1 General behaviour of the model

To test the general development of the ambiguity of the agent's grammar, I employ the same basic conditions that I have already used to study the development of expressivity in the model. These conditions are repeated in (18) below. In this setup, a simple semantic space made up of four conceptual units (A, B, X, Y), two of which are producible as signals and thus constitute the phonological space, is used. Meanings contain a maximum of two semantic units, and individual contexts provide one inferable or ignorable semantic unit. In this basic setup, no loss occurs, and the agent chooses randomly if he faces a situation of synonymy.

(18)	Semantic units:	A, B, X, Y
	Maximal meaning complexity:	2 semantic units
	Initial grammar:	$X \rightarrow X$
		$Y \rightarrow Y$
	Maximal context size:	1 semantic unit
	Entrenchment threshold:	$t = 0.0$
	Selection strategy:	random

Fig.5.21 documents the development of the ambiguity of the agent's grammar under these basic conditions. The dotted line represents the development of ambiguity in an individual simulation run, the solid line shows the average over 100 independent simulation runs. The latter is plotted against the corresponding average development of expressivity in Fig. 5.22, where the horizontal dashed line indicates the maximal level of expressivity, and the vertical dashed line indicates the point at which the agent's grammar reaches this level. The two figures show that ambiguity initially rises relatively steeply while the expressivity of

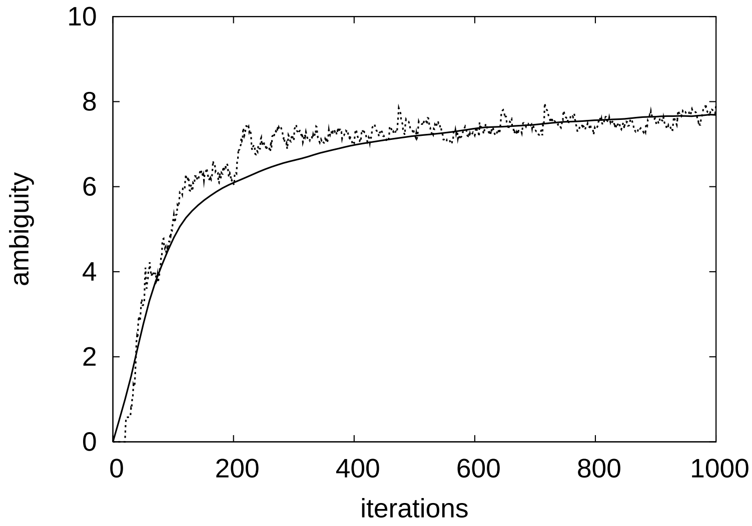


Figure 5.21: Development of the ambiguity of an agent's grammar under the idealised conditions listed in (18). The solid line represents the average ambiguity over 100 independent runs. The dotted line represents an individual run.

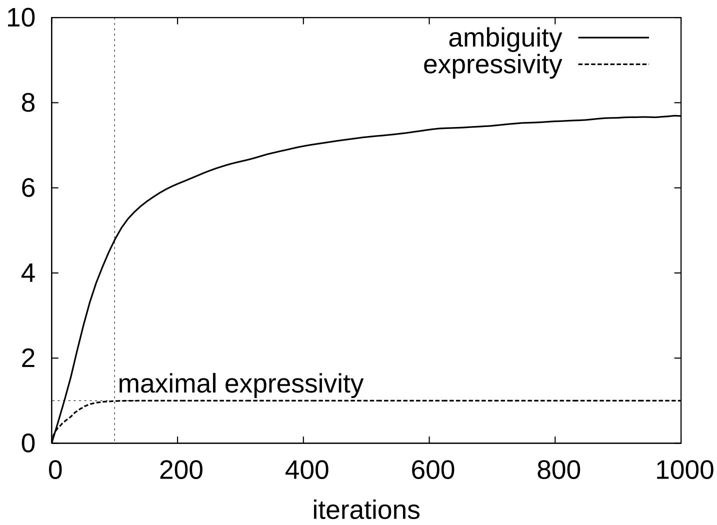


Figure 5.22: Development of expressivity and ambiguity under idealised conditions. Both lines represent the average of the respective value over 100 simulation runs under the conditions given in (18). The horizontal dashed line indicates the maximal level of expressivity, the vertical dashed line indicates the point at which the agent's grammar reaches this level.

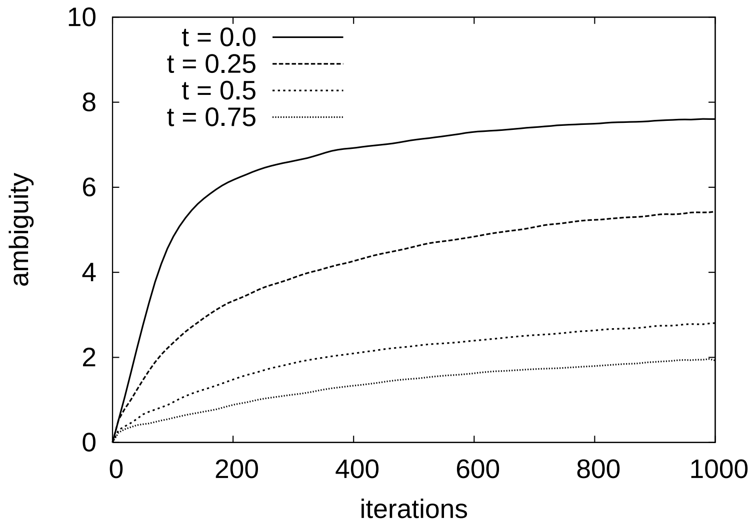


Figure 5.23: Development of ambiguity under different loss rates. Each line represents the average over 100 simulation runs. Apart from the entrenchment threshold *t*, which defines the rate of loss, all parameters are the same as in (18).

the agent's grammar is also increasing. Once the agent's grammar has reached maximal expressivity, its ambiguity continues to increase but to a more moderate degree. The reason why ambiguity continues to rise even after the grammar has become fully expressive has to be sought in the two facts that there is no loss and that the agent selects a signal at random if he has more than one option in a particular situation. The agent therefore does not necessarily choose extant conventions but continues to add ever new form-meaning associations to the grammar even though he would not need to if it simply was for communicative needs. Without loss, such new constructions continue to add ambiguity to the grammar. On the basis of this tentative explanation, we can thus formulate two hypotheses: (i) loss reduces ambiguity, and (ii) re-using extant conventions reduces ambiguity. These two hypotheses will be tested now.

5.4.2.2 Loss and re-use

The simulations confirm the first hypothesis: the introduction of loss reduces the degree of ambiguity that the agent's grammar reaches over the course of its cultural evolution. This observation is illustrated by Fig. 5.23. The figure plots the development of ambiguity under the same conditions as above but with different rates of loss. It is evident that in simulations where the agent's grammar faces higher rates of loss, the degree of ambiguity it reaches is lower than in simulations with lower loss rates. Loss reduces the amount of ambiguity that emerges in a grammar: the more constructions are lost in the course of the cultural evolution of a symbolic communication system, the lower is the degree of ambiguity which that system comes to exhibit.

The second hypothesis, which postulates that the re-using of extant forms reduces ambiguity, can be studied by varying the strategies that the agent applies for synonym selection. Fig. 5.24 plots the development of the ambiguity of the agent's grammar with various different strategies being employed. It shows that the level of ambiguity is by far the lowest if the agent consistently chooses the most entrenched signal. The other two consistent strategies (to choose the most underspecified signal and to choose the shortest signal) initially perform worse than if the agent selects a signal at random. However, after about 300 iterations, this trend is reversed and the level of ambiguity ceases to rise and stabilises at a comparatively moderate level where the agent applies the former strategies, whereas ambiguity continues to increase in the case where he chooses at random.

These observations comply with our hypothesis. An agent who consistently chooses the most entrenched form will introduce fewer new form-meaning associations to his grammar as he is biased towards re-using the ones that already exist. The other two consistent synonym selection strategies first have to establish a certain range of constructions, during which period ambiguity rises. But once such a range of form-meaning associations has been created, they can also start to re-use extant conventions. As an illustration of this two-stage process, imagine the following situation. An agent consistently chooses the most underspecified signal. In a first iteration, he possesses only a construction $X \rightarrow A$. However, he needs to communicate a meaning AB, and in the present context, B is inferable. The agent cannot but add a new construction $X \rightarrow AB$ to his code and thus increases its ambiguity. The next time the agent needs to convey AB in a situation where B is inferable, the agent will make use of the construction $X \rightarrow A$ again, even though a literal expression would now also be available. The agent chooses the most underspecified signal. However, this time, no new construction $X \rightarrow AB$ is added to the grammar anymore: after all, this construction already exists. Ambiguity is therefore not increased any further.

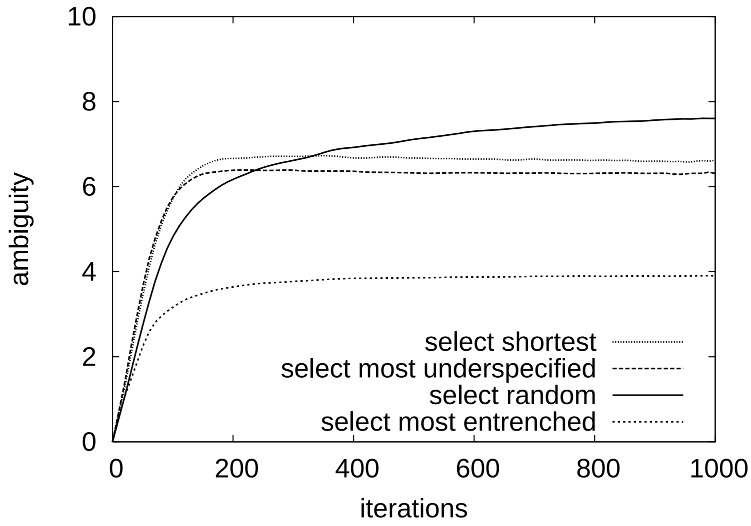


Figure 5.24: Development of ambiguity with different synonym-selection strategies and no loss. Each line represents the average over 100 simulation runs. All parameters other than the used synonym-selection strategy are as in (18).

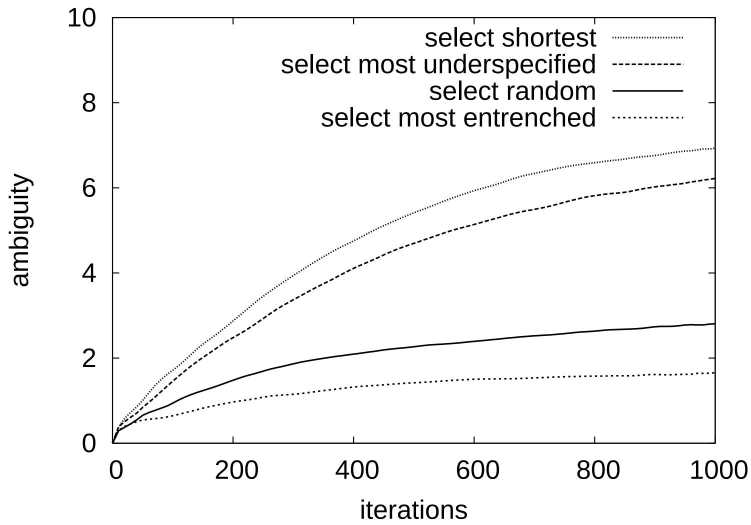


Figure 5.25: Development of ambiguity with different synonym-selection strategies and loss. Each line represents the average over 100 simulation runs. The entrenchment threshold t, which defines the rate of loss, is set to 0.5. All parameters other than the used synonym-selection strategy and the entrenchment threshold are as in (18)

The fact that consistent synonym-selection strategies other than the one that always chooses the most entrenched signal depend on a certain base of established form-meaning associations for ambiguity to stabilise suggests that loss would be detrimental, as it would constantly damage that base. It would be much more difficult for the agent's grammar to reach a stage where no new ambiguity would be added. This hypothesis is corroborated by Fig. 5.25. The figure plots the development of the ambiguity of the agent's grammar under the same conditions as in Fig. 5.24 before, with the one exception that now, loss has been introduced by setting the entrenchment threshold t to 0.5. While ambiguity continues to be low if the agent consistently picks the most entrenched signals, the other two consistent synonym-selection strategies now lead to degrees of ambiguity that are considerably higher than that yielded under the condition of random selection.

In summary, the presented simulation results corroborate the two initially formulated hypotheses that loss and the re-use of extant conventions each reduce the degree of ambiguity that emerges in an evolving symbolic communication system. The condition under which ambiguity is lowest is one in which loss is high and the agent consistently chooses the most entrenched signal. However, these conditions have previously proven to be detrimental to the development of the expressivity of a grammar (cf. section 5.2.2.2). We must therefore conclude that symbolic communication systems come to exhibit a certain trade-off between these two aspects of design: they become particularly expressive if the conditions in which they evolve allow for an increase in expressivity that is high but not high enough for the resulting degree of ambiguity to become dysfunctional.[16]

That human language indeed exhibits such trade-offs between competing functional pressures has been extensively documented, and especially typologists invoke such competing motivations as an explanation for language change and cross-linguistic variation (e.g. Langacker 1977; Givón 1979; Haiman 1985; DuBois 1987; Hall 1992; Kirby 1997; Croft 2000; Hopper and Traugott 2003). Langacker (1977:102), for instance, postulates that "[l]anguage change reflects the pressure to achieve linguistic optimality, but linguistic optimality has numerous dimensions reflecting the multi-faceted character of language, and the tendencies to

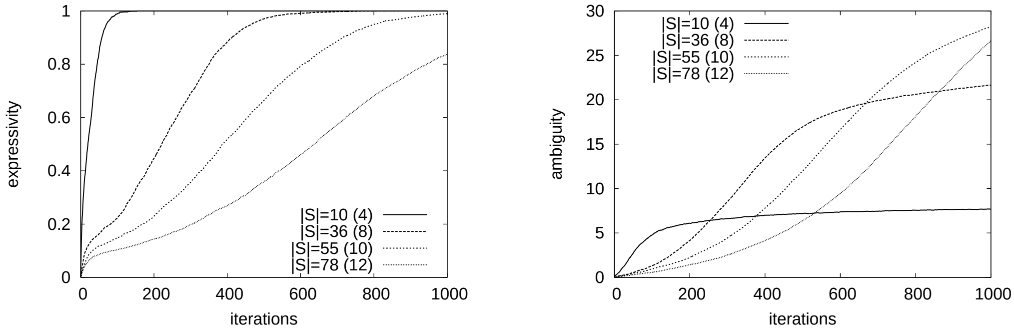
[16] At the same time, however, it needs to be emphasised again that linguistic evidence suggests that even relatively high levels of ambiguity are still not dysfunctional for communication: refer again to Traugott and Dasher's (2005) statement, which I have cited in the introduction to section 5.4.

achieve these different kinds of optimality are often in opposition to one another." Langacker (1977:128f.) observes that signal simplicity conflicts with expressivity (to which he refers as "perceptual optimality"), and that transparency (the absence of ambiguity and synonymy) conflicts with both code simplicity and signal simplicity—the latter being consistent with my argument that ambiguity allows for a higher degree of signal economy. The presented simulation has made another such conflict manifest: the conflict between transparency and expressivity. Note, however, that the discussed computer simulations show that the emergence of a trade-off between two competing pressures can be much less purpose-driven than the typologists' "competing-motivations" explanations would appear to suggest: in the simulations, a trade-off between a maximisation of expressivity and a minimisation of ambiguity can simply emerge due to environmental conditions of cultural evolution (the rate of loss and the employed synonym selection strategy)—the agent himself does not (and in fact cannot) "pull" toward either more expressivity or less ambiguity.

5.4.2.3 Ambiguity vs. expressivity?

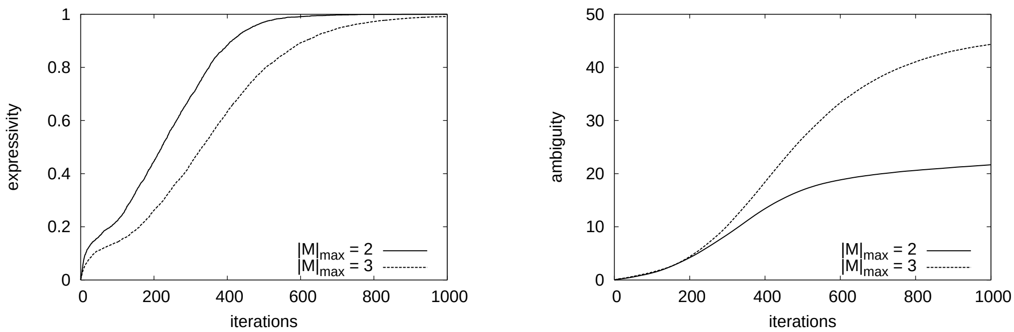
In the last part of this section, I explore if other parameters of the model yield similar conflicts between expressivity and ambiguity. To this aim, I employ some of the test conditions that I have already used in section 5.2.2 to study the development of expressivity. This will allow me to compare the impact that the respective factors have on expressivity with their impact on ambiguity. Figs. 5.26–5.29 compare the effect that the size of the semantic space, the complexity of meanings, the size of the context and the probability structure of the semantic space have on the development of the expressivity of a grammar (shown in the respective left-hand side sub-figure) with their effect on its ambiguity (shown in the respective right-hand side sub-figure).

It can be observed that a bigger semantic space (Fig. 5.26) is disadvantageous to both the development of the expressivity as well as the degree of ambiguity of a grammar: expressivity rises more slowly while ambiguity stabilises at a higher level if the size of the semantic space is increased. The same negative effect on both expressivity and ambiguity can be observed if the complexity of the meanings is increased (Fig. 5.27). These two parameters thus do not produce any tension between the pressure for increased expressivity and the ideal of low ambiguity: both come about if the semantic space is small and meanings relatively simple.



(a) Expressivity (b) Ambiguity

Figure 5.26: Impact of the size of the semantic space on expressivity and ambiguity. Each line represents the average over 100 simulation runs. All parameters other than the size of the semantic space are as in (18). $|S|$ indicates the number of meanings in the semantic space. The number of elementary semantic units is given in parentheses.



(a) Expressivity (b) Ambiguity

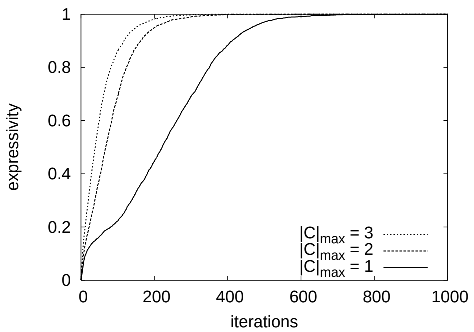
Figure 5.27: Impact of the meaning complexity on expressivity and ambiguity. Each line represents the average over 100 simulation runs. All parameters other than the maximal meaning complexity and the size of the semantic space, which is set to 36, are as in (18).

The case is different when it comes to the size of the context (i.e. the amount of inferable and ignorable information available in individual usage situations). Fig. 5.28(a) shows that the expressivity of the agent's grammar rises faster the more contextual material he can make use of when producing a signal. Fig. 5.28(b), on the other hand, illustrates that the more information the context contains, the more ambiguous the emerging grammar is going to be. Here, we therefore do have a conflict between the two pressures: with more inferences being drawn, codes end up more expressive but also more ambiguous.

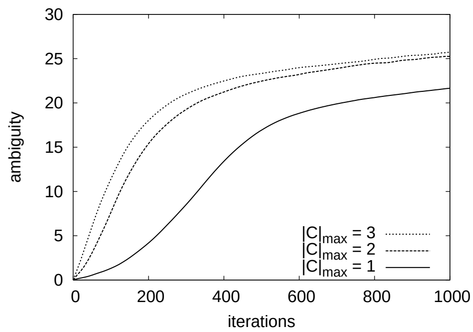
A final aspect that needs to be looked at is the probability structure of the semantic space. In section 5.2.2.2, I illustrated how non-flat probability distributions in the semantic space can have a substantial impact on the development of the expressivity of a grammar. Fig. 5.29(a) repeats this finding. It shows that of the four patterns introduced in Table 5.1, only one (pattern 1) is detrimental to the development of expressivity. In contrast, Fig. 5.29(b) shows that with regard to ambiguity, it suffices to have *any* sort of non-flat probability distribution: all four investigated patterns result in lower degrees of ambiguity than a semantic space with a flat structure. Pronounced probability distributions in the semantic space lead to reduced ambiguity in the grammar.

5.4.2.4 Summary and interpretation

The discussed computer simulations have been employed to study if there are factors other than phonological change and conscious avoidance of ambiguity that keep the degree of ambiguity that a grammar comes to exhibit in the course of its cultural evolution low and thus help to reduce the risk of massive ambiguity reaching a level where it can obstruct communication. We have seen that ambiguity is lower if (i) there is loss, (ii) if the agent uses a consistent synonym-selection strategy, in particular if he always chooses the most entrenched signal, (iii) if the semantic space is small, (iv) if meanings have low complexity, (v) if only little contextual information is available in specific usage situations, and finally (vi) if the semantic space exhibits a non-flat probability structure. In summary, we can say that the conditions under which the cultural evolution of a symbolic communication system happens also influence the degree to which it comes to exhibit ambiguity. It appears very likely that the cultural evolution of human language has been accompanied by both loss and a bias of its users to re-use the most entrenched constructions. The discussed simulations suggest that both these factors have kept the amount of ambiguity at a manageable level.

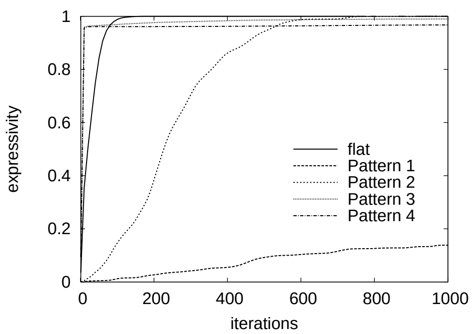


(a) Expressivity

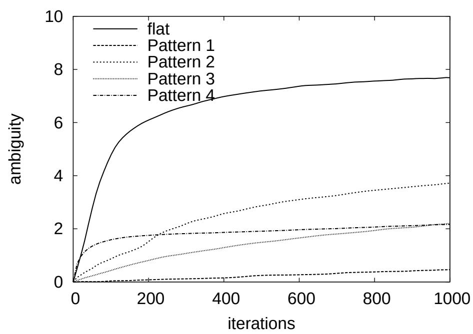


(b) Ambiguity

Figure 5.28: Impact of the context size on expressivity and ambiguity. Each line represents the average over 100 simulation runs. All parameters other than the context size and the size of the semantic space, which is set to 36, are as in (18).



(a) Expressivity



(b) Ambiguity

Figure 5.29: Impact of the structure of the semantic space on expressivity and ambiguity. Each line represents the average over 100 simulation runs. The used probability patterns are listed in Table 5.1. All other parameters are as in (18).

5.5 Conclusion

In this chapter, I have addressed the design puzzle: how does language come to exhibit the appearance of design for communication? In particular, I have looked at (i) how a symbolic communication system adapts its *expressivity* to the communicative needs of its users, and (ii) how it accommodates the fact that articulation constitutes a bottleneck in the communicative process by maintaining *signal economy*. The theoretical arguments and computer simulations provided have shown that a model of the cultural evolution of language that incorporates the concept of pragmatic plasticity is capable of accounting for how these two aspects of design for communication emerge over time.

Specifically, I have argued that pragmatic plasticity (in combination with conventionalisation) equips the users of symbolic communication systems with "tools" (i) to enhance the expressivity of those systems, and (ii) to perform context-specific data-compression on their signals. With the help of computer simulations, I have demonstrated how, and under which conditions, the iterated employment of pragmatic plasticity in individual usage situations eventually leads to conventionalised communication systems that are adapted both to the communicative needs of their users and to the articulation bottleneck. I conclude that the same set of cognitive mechanisms that can give rise to symbolism and grammar (cf. chapter 3) also lead, in the course of cultural evolution, to the emergence of core aspects of design for communication.

I have also addressed the role that ambiguity, a feature that is often characterised as dysfunctional, plays with regard to the design puzzle. I have argued that, contrary to the customary view, the availability of ambiguity in a symbolic communication system is not only not dysfunctional but actually contributes to a maximisation of the two aforementioned aspects of design for communication: expressivity and signal economy. In the form of layering, ambiguity is necessary for pragmatic plasticity to unfold its expressivity-enhancing power, and the availability of ambiguity allows codes to reach levels of signal economy that could not be achieved otherwise. It is only the *degree* of ambiguity in an individual grammar, not ambiguity as a feature itself, that can potentially become an impediment for communication. The computer simulations employed here have shown that certain environmental conditions keep the degree of ambiguity that a grammar develops in the course of its cultural evolution low and thus reduce the risk of it reaching a level at which it becomes dysfunctional.

In summary, the simulation results suggest that for a conventionalised communication system to become as expressive as human language, the individuals that develop it must have good memory capacities, so that little of what they observe is ever lost, and at the same time, they need to be able to make use of extensive amounts of contextual information. On the other hand, it seems that the presence of mechanisms of automatisation (in the form of re-using entrenched signal-meaning pairs) contributes to an adaptation of the system to the articulation bottleneck and to keeping its ambiguity at a level where it does not constitute an impediment for communication. The availability of refined capacities of both memorisation and automatisation may therefore be a part of the explanation of why humans have language but other animals do not.

CHAPTER 6

Conclusions

In the absence of observable evidence, how can we formulate theories of the evolution of language?[1] I have based this thesis on the following answer: we must assume the uniformitarian principle and postulate that the processes at work in the past are the same that can be observed in the present. The mechanisms invoked to account for the emergence and initial evolution of language must thus be consistent with those of present-day language use and change. Additionally, the enterprise to explain language as a cultural entity that is based on domain-general cognitive mechanisms suggests that the processes at work in the evolution of language are not in principle different from those identified for general cultural evolution. The account of language evolution developed in this thesis has therefore aimed to be consistent with models of cultural evolution in general and with models of language use and change in particular.

My discussion of such models has shown that there are *two* loci at which innovation can be introduced in the course of cultural evolution: use and cultural transmission. It can and must be assumed that both have played their part in the evolution of language. However, contemporary studies have mainly focused on the role of imperfect cultural transmission. To isolate the effect of this locus of innovation, they have consequently assumed an idealised representation of language use. In particular, they have failed to recognise the distinction between signal meaning and speaker meaning and ignored the fact that language use exhibits pragmatic plasticity. I have argued that such models of language evolution, while they have provided compelling explanations for the emergence of specific linguistic properties, fail to account for two basic evolutionary puzzles:

[1] A short version of this chapter appears as Hoefler (2008).

they cannot—and also do not claim to—explain (i) how language has emerged from no language and (ii) how it has come to exhibit the appearance of design for communication, which I have labelled the "emergence puzzle" and the "design puzzle" respectively.

To account for these two puzzles, my thesis has turned its focus to the role that innovative use plays in linguistic cultural evolution. Specifically, I have aimed (i) to develop a mechanistic model of the cultural evolution of language that acknowledges and incorporates the fact that language use exhibits pragmatic plasticity and (ii) to explore the explanatory potential of this fact with regard to the emergence puzzle and the design puzzle. In what follows, I summarise the contributions that the thesis has made to these two objectives, assess the significance of its findings for the study of language evolution and point to the possibilities it offers for future research.

6.1 Key elements of the proposed model

The mechanistic model of the cultural evolution of language that I have both introduced conceptually and implemented computationally is based on the following key elements.

- *Innovative use.* I have argued that innovations come about when language use exhibits pragmatic plasticity, that is, when, in a specific context, a signal comes to communicate a speaker meaning that differs from its conventional signal meaning. I have modelled the relationship between signal meaning and speaker meaning in terms of underspecification and overspecification.

- *Faithful learning.* To isolate the effect of innovative use, a maximally simple form of learning has been assumed: the mere storage of observed co-occurrences of produced forms and communicated meanings. This form of exemplar-based learning is psychologically particularly plausible and offers a way of modelling the process of conventionalisation through usage-memorisation and entrenchment.

Conceptually, the introduced model is built on a number of original arguments that call for the dissolution of traditionally assumed dichotomies and for the recognition of a number of continua:

- *Code-context continuum.* Code has been analysed as a special type of common-ground knowledge contained in the context of an utterance: the knowledge of a conventional association of a producible form with a specific meaning.

- *Coding-inference continuum.* Consequently, the process of coding has been re-analysed as a special case of inference from context that is only different from other forms of inference from context by its making use of the aforementioned special type of common-ground knowledge.

- *Form-meaning continuum.* Furthermore, I have pointed out that forms, because they are conceptual too, can themselves serve as signal meanings if they are not conventionally associated with any other meaning yet. Forms represent a special type of meanings, namely meanings for which direct evidence in form of a producible signal can be provided. I have shown that the recognition of this fact allows for the modelling of iconicity as well as the simulation of the transition from non-symbolic communication to symbolic communication.

- *Semantics-pragmatics continuum.* In the introduced model, linguistic forms acquire conventional (semantic) meaning gradually when users remember and further entrench their association with the meanings whose inference from context they have triggered. There is thus no clear-cut borderline between pragmatically inferred and semantically encoded meaning—even less so if one also takes the postulated code-context and coding-inference continua into account.

The introduced model also recognises the so-called symbolic thesis postulated by construction-based approaches to grammar, which dissolves the dichotomy between lexicon and syntax:

- *Syntax-lexicon continuum.* Syntactic constructions are conceived as symbolic associations between a form and meaning. In this sense, they are not different from lexical items. The two only differ in their degree of specificity, with syntactic constructions being more schematic than lexical ones. Because syntactic constructions are conventional form-meaning associations, they can also exhibit pragmatic plasticity in specific contexts of use.

6.2 The explanatory potential of pragmatic plasticity

The key contribution of this thesis to the study of language evolution is its exploration of the explanatory potential of pragmatic plasticity with regard to the emergence puzzle and the design puzzle.

6.2.1 Explaining the emergence puzzle

The thesis has shown that a model of the cultural evolution of language that incorporates the notion of pragmatic plasticity has the potential to account for the following aspects of the emergence puzzle:

- *The link between linguistic and non-linguistic communication.* Pragmatic plasticity points to the ostensive-inferential nature of linguistic communication and therefore to the origin of language in this pre-symbolic form of communication. This thesis has bridged the evolutionary gap between no language and language by identifying ostensive-inferential communication as the continual aspect that is present in both stages, and suggested that the cognitive mechanisms involved in ostensive-inferential communication may be sufficient for the transition from one stage to the other.

- *The emergence of symbolism.* In particular, it has shown how the pragmatic plasticity exhibited by pre-symbolic communicative cues, and the conventionalisation of the usages of such cues, can lead to the emergence of symbolism. Through iterated usage-induced semantic change, the relationship between a form and its conventional meaning can become obscured, and eventually appear arbitrary. Such semantic change is initiated by under- and/or overspecified re-use, that is, when extant conventions come to exhibit pragmatic plasticity in specific contexts of use.

- *The emergence of grammar.* I have sketched how the same set of cognitive mechanisms can also be applied to account for the processes of syntacicisation and grammaticalisation. Firstly, I suggested that pragmatic plasticity can explain why abstract schemata, for instance the order in which two symbols are produced, become conventionalised: the schemata themselves serve as communicative cues that lead to the inference of meaning from context. Secondly, I argued that through pragmatic plasticity in re-use, extant conventions, both specific and schematic, can undergo further semantic change and eventually become more "functional."

- *The protolanguage issue.* I have consequently argued that the assumption of two distinct stages in the evolution of language, one with symbolism but without grammar ("protolanguage") and one with both symbolism and grammar, would not be warranted if the same set of cognitive mechanisms were to lead to both the emergence of symbolism and the emergence of grammar. A model that incorporates pragmatic plasticity would then present a scenario of language evolution in which both symbolism and grammar gradually co-evolve from iterated ostensive-inferential communication.

6.2.2 *Explaining the design puzzle*

The thesis has also shown that a model that incorporates pragmatic plasticity has the potential to account for the presence of the following aspects of the appearance of design for communication in language.

- *Expressivity.* Pragmatic plasticity constitutes a "tool" to enhance the expressivity of a symbolic communication system and can therefore account for how language has come to adapt to the communicative needs of its user(s). Because language use exhibits pragmatic plasticity, extant conventions can be used as stepping stones to reach novel, previously inexpressible meanings. When such new usages are conventionalised, they can themselves serve as stepping stones in subsequent usage events. Pragmatic plasticity thus allows for the cumulative exploration of ever new meaning spaces. This process has been studied with the help of computer simulations in this thesis.

- *Signal economy.* In the form of under- and overspecification, pragmatic plasticity provides two "tools" for context-specific signal reduction and can thus account for how language has come to adapt to the fact that articulation constitutes a bottleneck that slows down the communicative process. I have shown with the help of computer simulations that under the right conditions the iterated use of under- and overspecification can help to keep the average signal length of a symbolic communication system low. Pragmatic plasticity enables language to function like an evolving lossy compression algorithm that adapts to its environment in the course of its cultural evolution.

- *Ambiguity.* Pragmatic plasticity also explains the ubiquity of the apparently dysfunctional feature of ambiguity in language. Because linguistic signals are interpreted in context anyway, ambiguous code does not constitute a problem for communication. However, ambiguity is even positively functional: a code that does not allow for ambiguity in the form of layering cannot make use of the expressivity-enhancing capacities of pragmatic plasticity. Furthermore, I have argued from an information-theoretic perspective that ambiguous codes allow for greater signal economy than non-ambiguous codes, which again adds to the appearance of design for communication.

6.3 Assessment and future research

As its main contribution to the study of language evolution, the presented thesis has suggested that a model of the cultural evolution of language that incorporates the notion of pragmatic plasticity has the capacity to account for (i) how language emerged from no language, and (ii) how language has come to exhibit the appearance of design for communication. With regard to explaining these two puzzles, a model that locates innovation in use and treats cultural transmission as faithful seems to be superior to one that locates innovation in imperfect cultural transmission and assumes use to be faithful.

This observation leads to a final hypothesis. Could it be that the mechanisms available to models that locate innovation in use are indeed *sufficient* for the emergence of language? This is not to suggest that the processes studied by transmission-based models did not occur in the course of language evolution but to open up the possibility that a complex symbolic communication system like language could have emerged and evolved just as well if they had not. If this was indeed the case, the following to factors would have played a crucial part in the emergence of langauge: (i) extended use of pragmatic plasticity, which in turn is based on a highly developed ability to recognise common ground, and (ii) an increased fidelity of learning, both individual and social.

It must be left to future research to test this hypothesis. The model that I have proposed provides (but also *only* provides) a framework in which such testing can be carried out. Among other things, the following questions remain to be addressed:

- Can phonological change be described in terms of under- and overspecification and thus a common cognitive basis for both phonological and semantic change be identified?
- Do the findings gained from the presented computer simulations scale up? That is, do they still hold if we run simulations that work with semantic and phonological spaces that are much larger than the ones assumed for the cases discussed?
- Can we simulate the emergence of complex syntactic phenomena by replacing the abstract forms and meanings used in the introduced computational model with more specific representations?
- What information can be gained from the computer simulations about communication systems potentially shifting from more to less iconicity in the course of their cultural evolution?
- Can the developed conceptual model be applied to describe experimental data such as that provided by Galantucci (2005) or Garrod et al. (2007)?

In summary, this thesis makes a case for the formulation of mechanistic models, both conceptual and computational, that describe the processes involved in the cultural evolution of language at a high level of abstraction. Such models allow us to gain an understanding of the mechanisms at work by generalising over established analyses of individual linguistic phenomena. After all, the origin of language cannot be explained at a level of description that is specific to language: rather, it needs to be accounted for in terms of more basic, domain-general mechanisms. If we want to be able to identify the minimal set of cognitive capacities that must have been in place for language to emerge, we must abstract away from merely linguistic descriptions and classifications and thereby separate the explanandum, language, from its explanans. The present thesis contributes to this goal.

APPENDIX A

Algorithms

This appendix gives a more thorough treatment of the algorithms that the computational model discussed in chapters 4 and 5 applies to generate speaker meanings (section A.1), to calculate the probabilities of individual meanings (section A.2), to generate the contextual information that is made available to the agent in each iteration (section A.3), and to produce a signal in light of a given speaker meaning and context (section A.4).

In the implementation, the semantic space is determined by two program parameters:

1. the set of all semantic units discernible by the agent
2. the maximal number of such units that a meaning can contain (henceforth referred to as the *maximal meaning length*)

Each semantic unit is assigned two weights: one indicates the probability by which the semantic unit occurs in a generated meaning and one the probability by which it occurs in the context.

A.1 Generating a meaning

In each iteration of the simulation, a meaning for the agent to communicate is generated by the algorithm performing the following steps:

1. Randomly choose a number between 1 and the pre-defined maximal meaning length. This number indicates the length of the meaning to be generated.

2. Randomly choose a semantic unit according to the probability distribution of the semantic units.
3. Add the chosen semantic unit to the meaning to be generated.
4. Randomly choose another semantic unit. If the chosen unit is already part of the meaning to be generated, discard it. If it is not, add it.
5. Repeat step 4 until the meaning has reached the length determined in step 1.

A.2 Calculating the probability of a meaning

Because of the way meanings are generated (see A.1 above), each permutation of a meaning can occur with a different probability. The probability that a meaning occurs is thus the sum of the probabilities of its permutations:

(1) $$P(M) = \sum_{i=1}^{|M|!} P(M_i) \text{ where } M_i \text{ is a permutation of } M$$

The probability that a particular permutation of a meaning occurs is captured by the following formula:

(2) $$P(M_i) = \frac{1}{l_{max}} \cdot \prod_{j=1}^{|M|} \frac{P(c_j)}{1 - \sum_{k=0}^{j-1} P(c_k)} \text{ where } c \in M_i$$

Formula (2) is best explained by working through an example. Let us assume that the experimenter has set the program parameters as given in (3):

(3)

Discernible semantic units:	A, B, C
Probability of occurrence in the speaker meaning:	$P(A) = 0.5$
	$P(B) = 0.25$
	$P(C) = 0.25$
Maximal meaning length:	3

The parameters in (3) yield the following semantic space:

(4) Semantic space: $\{ \{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}, \{A, B, C\} \}$

The probability that, for instance, meaning $\{A, B, C\}$ occurs is then the sum of the probabilities by which its permutations $[A, B, C]$, $[A, C, B]$, $[B, A, C]$, $[B, C, A]$,

$[C, A, B]$, $[C, B, A]$ occur. In correspondence with formula (2), the following procedure is applied to calculate the probability of an individual permutation, e.g. $[B, A, C]$:

1. Calculate the probability by which the given meaning length occurs (in our example, meanings can contain one, two or three units. The probability that a meaning has a length of 3 is therefore $\frac{1}{3}$).
2. For each unit contained in the permutation, calculate the probability by which it would have been selected at the respective position in the permutation:
   (a) Calculate the probability that B is selected first. (This is $P(B) = 0.25$.)
   (b) Calculate the probability that A is selected second. (Because B cannot be selected anymore, the choice is only between A and C now. The probability that A is chosen under these conditions is thus $\frac{P(A)}{1-P(B)} = \frac{0.5}{0.75}$).
   (c) Calculate the probability that C is selected third. (Because B and A cannot be selected anymore, the choice is down to C now. The probability that C is chosen under these conditions is thus $\frac{P(C)}{1-(P(B)+P(A))} = \frac{0.25}{0.25} = 1$).
3. The probability of the considered permutation is the product of the probability calculated in step 1 and all probabilities calculated in step 2. (In our example, this means that $P([B, A, C]) = \frac{1}{3} \cdot 0.25 \cdot \frac{0.5}{0.75} \cdot 1 = \frac{1}{32}$.)

To get the probability by which meaning $\{A, B, C\}$ occurs, the probabilities of all its permutations need to be summed up, as stated in formula (1). This yields $P(\{A, B, C\}) = \frac{1}{3}$ (which makes sense because only one meaning with three units is possible and each of the three possible meaning lengths is equally probable).

A.3 Generating a context

In each generation, a context is generated. This context can contain two types of information: semantic units that are inferable and semantic units that are ignorable. The former constitute semantic units that are part of the speaker meaning but do not need to be expressed explicitly because, in the present context, they can be inferred. The latter are semantic units that are *not* part of the speaker meaning but can nevertheless be part of the signal meaning since they constitute irrelevant information in the present context and will thus be ignored (cf.

section 4.1.2.1). The maximal number of semantic units that the context of an iteration can contain (henceforth referred to as the "maximal context size") needs to be pre-defined by the experimenter.

The algorithm that generates the contextual information functions as follows:

1. Randomly choose a number between 1 and the pre-defined maximal context size. This number indicates the number of semantic units to be generated for the given context.
2. Randomly choose a semantic unit according to the probability distribution of the semantic units.
3. If the chosen semantic unit already occurs in the generated speaker meaning, then mark it with a plus sign: it constitutes inferable information. If the the chosen semantic unit does not occur in the generated speaker meaning, mark it with a minus sign: it constitutes ignorable information.
4. Add the signed semantic unit to the context to be generated.
5. Randomly choose another semantic unit. If the chosen unit is already part of the context to be generated, discard it. If it is not, repeat steps 3 and 4.
6. Repeat step 5 until the context has reached the size determined in step 1.

A.4 Producing a signal

After a speaker meaning and a context have been generated, the agent applies his I-language to produce an appropriate signal. The process of signal production or use is described in informally in section 4.1.2. Here I provide the algorithm that the agent applies to produce a signal in the context of a given iteration. This algorithm consists of the following steps:

1. Generate all possible signal meanings for the given speaker meaning in the given context.
   A possible signal meaning P_i is defined as the union of the given speaker meaning S and a (potentially empty) subset I_i of all semantic units that are inferable in the given context, minus a (potentially empty) subset N_i of all semantic units that are ignorable in the given context:

   $P_i = (S \cup I_i) \setminus N_i$

Legend:	P_i	a possible signal meaning
	S	the given speaker meaning
	I	the given set of all inferable semantic units
	I_i	a (potentially empty) subset of I: $I_i \subseteq I$
	N	the given set of all ignorable semantic units
	N_i	a (potentially empty) subset of N: $N_i \subseteq N$

2. Generate all possible signals:
   (a) For each possible signal meaning, generate all signals that the agent's I-language maps it onto.
   (b) From the resulting set of signals exclude those that the agent's I-language also maps onto a meaning that contains semantic units that are neither part of the given speaker meaning nor constitute ignorable information in the given context. (This step is not implemented in the current model; see comments below.)
3. From the resulting list of signals, choose those signals that meet best the criteria given by the applied synonym selection strategy:
   (a) If the strategy is to choose the most entrenched signal, calculate for each signal the average entrenchment of the constructions that would be involved in its production and choose the signal(s) with the highest value.
   (b) If the strategy is to choose the most underspecified signal, count for each signal the number of semantic units that occur in the given speaker meaning but not in the signal meaning of that signal and choose the signal(s) with the highest value.
   (c) If the strategy is to choose the shortest signal, count the number of units contained in the signal and choose the signal(s) with the lowest value.
   (d) If the strategy is to choose a signal at random, do nothing.
4. If there is still more than one signal left, choose one at random.

Note that without step 2(b), which is not included in the current implementation of the model, the described algorithm can result in a situation of pragmatic ambiguity: the produced signal may be mapped to more than one signal meaning by the agent's grammar, and these signal meanings may identify different speaker meanings from the set of meanings that are potentially relevant in the given situation (cf. section 4.1.2.1).

As an example, imagine that in a given situation both *A* and *B* consitute relevant information, that the agent needs to communicate speaker meaning *A*, and that his grammar consists of the constructions $X \rightarrow A$ and $X \rightarrow B$. Without the addition of step 2(b), the algorithm described above would result in the agent producing signal *X*. From the perspective of the hearer (which is not modelled here), this would however lead to pragmatic plasticity: in the given context, *X* could denote *A* as well as *B*. The produced cue would thus not be sufficient for the hearer to work out which of the relevant meanings the agent wants to convey.

By adding step 2(b) to the algorithm, the occurrence of such a case can be prevented: the agent would now not be able to produce any signal for the intended speaker meaning under the described circumstances, and the simulation would move on to the next iteration (as described at the end of section 4.1.2.2). This function is, however, not included in the simulations discussed in the present thesis.

Appendix B

Measurements

This appendix provides detailled descriptions of the definitions and methods that have been used to measure expressivity, signal economy and ambiguity in the computer simulations discussed in chapter 5.

B.1 Measuring expressivity

I define the expressivity of a grammar as the proportion of the agent's semantic space that is covered by the conventionalised meanings which the grammar associates with forms. A preliminary way of measuring expressivity in the model is thus to divide the number of meanings expressible by the grammar by the number of possible meanings contained in the semantic space. This provisional definition of expressivity is summarised in formula (1) below—where *expr* stands for expressivity, G for the grammar, $S(G)$ for the semantic space covered by the grammar and $S(A)$ for the total semantic space available to the agent.

(1) $$expr(G) = \frac{|S(G)|}{|S(A)|}$$ *(provisional definition)*

In example (2) below, I apply this preliminary method of measuring expressivity. The example shows a grammar that covers two thirds of the semantic space: of the three meanings that the agent can possibly want to convey in a usage event (A, B, AB), two occur as conventionalised meanings in a construction contained in the grammar (A and B).

(2)	Semantic space:	*A*, *B*, *AB*
	Grammar:	$X \rightarrow A$
		$Y \rightarrow B$
	Expressivity:	2/3 = 0.67

In contrast, example (3) depicts a situation where the agent's grammar exhibits full expressivity because it contains a construction for every meaning in the semantic space. In summary, a grammar has maximal expressivity (expressivity = 1) if it covers the whole semantic space, and minimal or no expressivity (expressivity = 0) if it cannot express any meaning contained in the semantic space.

(3)	Semantic space:	*A*, *B*, *AB*
	Grammar:	$X \rightarrow A$
		$Y \rightarrow B$
		$XY \rightarrow AB$
	Expressivity:	3/3 = 1

However, formula (1) provides only a very sketchy description of expressivity. A more precise definition must also take into account that not all meanings contained in the semantic space necessarily occur with the same probability: some meanings may have to be communicated more often than others. In the computational model, this phenomenon can be simulated by assigning weights to the meanings contained in the semantic space of the agent which determine with what probability individual meanings are randomly selected to constitute the speaker meaning the agent needs to convey. This fact makes it necessary for expressivity to be calculated in a more sophisticated way.

Example (4) illustrates what impact the probabilities with which individual meanings occur must have on the expressivity measured in a grammar. Suppose a semantic space contains two meanings, *A* and *B*, and the former occurs with a probability of 0.9 as the agent's speaker meaning whereas the latter occurs only with a probability of 0.1. In such a semantic space, a grammar that contains both a construction for *A* and a construction for *B*, as shown in G_3 below, has maximal expressivity. The other limiting case is a grammar without any constructions at all, as represented by G_0. The expressivity value of such a grammar would be zero. The interesting distinction, however, is that between a grammar that only has a construction for *A* (like grammar G_1) and one that only has a construction for *B* (like grammar G_2). Even though both these grammars contain the same number of meanings, our intuitive notion is that G_1 is more expressive than G_2 because A occurs more frequently than B. We thus need a formula to measure expressivity that takes this goal into account.

(4)

Semantic space:	A, B
Probabilities:	$P(A) = 0.9$
	$P(B) = 0.1$
Grammar G_0:	–
Grammar G_1:	$X \rightarrow A$
Grammar G_2:	$Y \rightarrow B$
Grammar G_3:	$X \rightarrow A$
	$Y \rightarrow B$
Limiting cases:	$expr(G_0) = 0$
	$expr(G_3) = 1$
Goal:	$expr(G_1) > expr(G_2)$

I therefore provide a definition of expressivity that weights each meaning contained in the grammar according to its probability of occurrence. This goal is realised in formula (5), which calculates the expressivity of a grammar G as the sum of the probabilities P of all meanings M contained in the semantic space $S(G)$ covered by the grammar.[1]

(5) $$expr(G) = \sum_{i=1}^{|S(G)|} P(M_i) \text{ where } M \in S(G)$$ *(final definition)*

According to formula (5), the expressivity of grammar G_1 in (4) is 0.9, whereas the expressivity of grammar G_2 is only 0.1.

B.2 Measuring signal economy

To measure a grammar's adaptation to the articulation bottleneck, I have adapted an information-theoretic method to the specifics of the computational model. In the domain of data compression, the degree to which a code exhibits signal economy is determined by calculating the *average codeword length* of that code.[2] The term "codeword" denotes the form side of a form-meaning mapping.

[1]Refer to appendix A.2 for a detailed account of how the probability of occurrence of a meaning is determined in the computational model.

[2]The same concept is at times also referred to as the "average length" of the code or the "rate" of the code (cf. Sayood 2006:6, 28).

The concept of average codeword length can be applied to the introduced computational model: it corresponds to the average length of the meaning-bearing forms in the agent's grammar, which I have referred to as the grammar's *average signal length* in chapter 5. The simplest definition of the average signal length of a grammar is given in formula (6).

(6) $$asl(G) = \frac{\sum_{i=1}^{|G|} |F_i|}{|G|} \text{ where } F_i \rightarrow M_i \in G$$ *(first provisional definition)*

By this definition, the average signal length *asl* of a grammar G is the sum of the lengths of all forms F that occur in the grammar, divided by the number of forms (or rather the number of all form-meaning mappings $F \rightarrow M$) contained in the grammar. The average signal length of the grammar in (7) would thus be $\frac{1+2}{2} = 1.5$.

(7) $X \rightarrow A$
$YZ \rightarrow B$

However, definition (6) does not take into consideration the fact that not all meanings necessarily occur with equal probability. Consider, for instance, the two grammars given in example (8).

(8)		
	Probabilities:	$P(A) = 0.9$
		$P(B) = 0.1$
	Grammar G_1:	$X \rightarrow A$
		$YZ \rightarrow B$
	Grammar G_2:	$XZ \rightarrow A$
		$Y \rightarrow B$
	Goal:	$asl(G_1) < asl(G_2)$

Both grammar G_1 and grammar G_2 have a construction for meaning A and a construction for meaning B. If formula (6) were to be applied, both grammars would exhibit the same average signal length since both contain one construction whose form consists of one semantic unit and one construction whose form is made up of two semantic units. However, example (8) also provides the probability of occurrence of the meanings encoded by the grammar (refer to appendix A.2 for a detailed account of how the probability of occurrence of a meaning can be determined in the model): meaning A is very frequent whereas meaning B is rare. In this situation, it makes sense to say that grammar G_1, where the signal for the more frequent meaning is shorter than that for the less frequent one, has a lower average signal length than grammar G_2, where the signal for the more frequent meaning is longer than that for the less frequent one. This goal is taken into account in the revised formula shown in (9), which defines the average signal length of a grammar as the sum of the length of each form, times the probability of occurrence of the meaning associated with that form.

(9) $$asl(G) = \sum_{i=1}^{|G|} P(M_i)|F_i| \text{ where } F_i \rightarrow M_i \in G$$ *(second provisional definition)*

If it is determined according to formula (9), the average signal length of grammar G_1 in example (8) amounts to 1.1, whereas the average signal length of grammar G_2 is 1.9. The revised definition thus meets our goal.

Further complication is introduced by the fact that not all grammars exhibit maximal expressivity. If a grammar is not fully expressive, its average signal length would be distorted if it was calculated according to definition (9). In example (10), for instance, grammar G_3 covers only two out of three possible meanings, and its average signal length would consequently appear to be 0.75 instead of 1.5.

(10)		
	Probabilities:	$P(A) = 0.25$
		$P(B) = 0.25$
		$P(AB) = 0.5$
	Grammar G_3:	$X \rightarrow A$
		$YZ \rightarrow B$
	Goal:	$asl(G_3) = 1.5$

To fix this problem, one must modify the definition of the average signal length of a grammar as given in formula (11). This formula divides the sum of individual signal lengths and meaning probabilities by the total expressivity of the grammar.

(11) $$asl(G) = \frac{\sum_{i=1}^{|G|} P(M_i)|F_i|}{expr(G)} \text{ where } F_i \rightarrow M_i \in G$$ *(third provisional definition)*

A final issue that needs to be addressed is exemplified in (12): how do we deal with synonymous form-meaning mappings, that is, with constructions that express the same meaning? In grammar G_4, meaning AB constitutes the right-hand side of two constructions, one mapping it onto form XY, the other one mapping it onto form Z.

(12) Probabilities: $P(A) = 0.25$
$P(B) = 0.25$
$P(C) = 0.5$
Grammar G_4: $X \rightarrow A$
$Y \rightarrow B$
$XY \rightarrow AB$
$Z \rightarrow AB$

Goal: $asl(G_4) = 1.25$

If formula (11) were to be applied to determine the average signal length of this grammar, the result would again be distorted. As a solution, I therefore devise the final definition of average signal length given in (13) below. To prevent the described problem, each time the probability of the meaning M of a construction is used, it is divided by the number n of constructions in grammar G whose right-hand side that meaning constitutes.

(13) $asl(G) = \frac{\sum_{i=1}^{|G|} \frac{P(M_i)}{n(M_i)} |F_i|}{expr(G)}$ where $F_i \rightarrow M_i \in G$ *(final definition)*

B.3 Measuring ambiguity

In a first instance, the ambiguity of a grammar can be defined as the number of ambiguous forms (i.e. the number of forms associated with more than one meaning) which that grammar contains. According to this notion of ambiguity, grammar G_1 in (14) contains one instance of ambiguity (the form X) and grammar G_2 two (the forms X and Z):

(14) Grammar G_1: $X \rightarrow A$
$X \rightarrow B$
$Y \rightarrow C$
$Z \rightarrow D$

Grammar G_2: $X \rightarrow A$
$X \rightarrow B$
$Y \rightarrow C$
$Z \rightarrow D$
$Z \rightarrow E$

Goal: $amb(G_1) < amb(G_2)$

Consider now the two grammars in (15). Both grammars contain just one ambigous form, X. According to the described simple notion of ambiguity they would thus be equally ambiguous. However, the fact a definition of ambiguity needs to capture in this case is that a form with three meanings is more ambiguous than a form with two meanings. To this aim, I devise the revised definition of ambiguity given in (16).

(15) Grammar G_1: $X \rightarrow A$
$X \rightarrow B$

Grammar G_2: $X \rightarrow A$
$X \rightarrow B$
$X \rightarrow C$

Goal: $amb(G_1) < amb(G_2)$

(16) a. $amb(G) = \sum amb(F)$

b. $amb(F) = |C_F| - 1$ *(provisional definition)*

Legend:

G	a grammar
F	a form occuring in G
C_F	the set of all constructions $x \in G$ with form F

Definition (16a) describes the ambiguity of a grammar G as the sum of the ambiguity of all forms F occuring in G. In (16b), the ambiguity of a form F is then defined as the number of constructions in G that have that form, minus 1. By subtracting 1 from the number of constructions with the same form, we achieve a more intuitive notion of ambiguity: a form with only one meaning has ambiguiy 0, a form with two meanings has ambiguity 1, a form with three meanings has ambiguity 2, and so on.

We also have to consider the fact that form-meaning associations can be entrenched to different degrees. To this aim, compare the two grammars in (17), in which the relative entrenchment of each construction is given in parentheses.[3]

(17)	Grammar G_1:	$X \rightarrow A$	(0.5)
		$X \rightarrow B$	(0.5)
	Grammar G_2:	$X \rightarrow A$	(0.9)
		$X \rightarrow B$	(0.1)
	Goal:	$amb(G_1) > amb(G_2)$	

Both grammars in (17) consist of just one form that is associated with two meanings. According to definition (16), they would thus exhibit the same degree of ambiguity. However, there is a difference in the two grammars with regard to the ease with which the two meanings of the respective form can or will be recovered. In grammar G_1, both meanings of form X are equally entrenched and thus equally accessible. This is not the case for grammar G_2. Here, the use of form F is much more likely to elicit meaning A than meaning B because the assocation between F and A is much more entrenched than the assocation between F and B. Grammar G_1 is thus more ambiguous than grammar G_2. This fact is captured in the revised definition of ambiguity presented in (18).

(18) $$amb(F) = \frac{|C_F| - 1}{1 + \sum_{i=1}^{|C_F|} \Delta(c_i)}$$ *(provisional definition)*

[3] The relative entrenchment of a construction is obtained by dividing its absolute entrenchment by the sum of the entrenchment of all constructions occurring in the respective grammar. Like this, a normalisation of the entrenchment values of the individual constructions in a grammar can be achieved.

Legend:	G	a grammar
	F	a form occuring in G
	C_F	the set of all constructions $x \in G$ with form F
	c	a construction $\in C_F$
	$\Delta(c)$	the deviation of the relative entrenchment of c from the average relative entrenchment of C_F

Definition (18) makes sure that constructions with the same form whose entrenchment differs a lot are considered less ambiguous than constructions with the same form whose entrenchment is very similar. To this end, the deviation of the entrenchment of each construction from the average entrenchment of all constructions is calculated.[4] The ambiguity value determined by formula (16b) is then divided by the sum of all these deviations. To prevent division by zero, this sum is increased by 1.

In grammar G_1 of example (17), the deviation of the entrenchment of either construction from the average entrenchment of the two constructions is 0. The earlier ambiguity value 1 is thus divided by 1: the ambiguity of G_1 remains 1. In constrast, the entrenchment of either construction in grammar G_2 deviates by 0.4 from the average entrenchment (0.5). The original ambiguity value 1 is thus divided by $1 + \Delta(X) = 1.8$ (i.e. $1 + 0.4 + 0.4$), and definition (18) yields an ambiguity value of only 0.56 for the form X in grammar G_2. As intended, the revised definition of ambiguity acknowledges that grammar G_2 is less ambiguous than grammar G_1.

The last shortcoming of our definition of the ambiguity of a grammar is that it does not take into account how many non-ambiguous forms the grammar contains. According to the formula devised so far, both grammars in example (19) are equally ambiguous, even though grammar G_1 consists of only two constructions with the same form, whereas grammar G_2 additionally contains two (much more entrenched) non-ambiguous constructions. An ideal definition of ambiguity would conceive G_1 as more ambiguous than G_2.

[4] Remember that the *relative* entrenchment of the constructions is used: the sum of the entrenchment of all constructions in a grammar always amounts to 1.

(19) Grammar G_1: $X \rightarrow A$ (0.5)
$X \rightarrow B$ (0.5)

Grammar G_2: $X \rightarrow A$ (0.1)
$X \rightarrow B$ (0.1)
$Y \rightarrow C$ (0.4)
$Z \rightarrow D$ (0.4)

Goal: $amb(G_1) > amb(G_2)$

To ensure that grammar G_1 is rated as more ambiguous than grammar G_2, the calculated ambiguity value must be put in relation to the proportion of the grammar that is ambiguous in the first place. This proportion can be obtained by adding up the relative entrenchment of all constructions in the grammar whose forms are associated with more than one meaning. I therefore devise the final definition of ambiguity stated in (20).

(20) a. $amb(G) = \sum amb(F)$

b. $amb(F) = \dfrac{|C_F| - 1}{1 + \sum_{i=1}^{|C_F|} \Delta(c_i)} \cdot \sum_{j=1}^{|C_F|} R(c_j)$ *(final definition)*

Legend:		
	G	a grammar
	F	a form occuring in G
	C_F	the set of all constructions $x \in G$ with form F
	c	a construction $\in C_F$
	$\Delta(c)$	the deviation of the relative entrenchment of c from the average relative entrenchment of C_F
	$R(c)$	the relative entrenchment of construction c in grammar G

Formula (20) multiplies the ambiguity value calculated for a specific form F with the sum of the relative entrenchments R of all constructions c containing that form. In grammar G_1 of example (19) above, the ambiguity value for X is thus multiplied by a factor 1 (i.e. the sum of 0.5 and 0.5). The resulting ambiguity of grammar G_1 is consequently 1. In grammar G_2, however, the calculated ambiguity value of form X is multiplied by a factor 0.2 (i.e. the sum of 0.1 and 0.1). The ambiguity of grammar G_2 therefore only amounts to 0.2. The definition of ambiguity that I have devised thus acknowledges that grammar G_2 is less ambiguous than grammar G_1.

Appendix C

Publications

This appendix contains three articles published prior to the completion of this thesis:

Hoefler, S. (2008). Pragmatic plasticity: A pivotal design feature? In Smith, A. D. M., Smith, K., and Ferrer i Cancho, R., editors, *The Evolution of Language: Proceedings of the 7th International Conference on the Evolution of Language*, pages 427–428, Barcelona. World Scientific Press.

Hoefler, S., and Smith, A. D. M. (2008). Reanalysis vs. metaphor? What grammaticalisation *can* tell us about language evolution. In Smith, A. D. M., Smith, K., and Ferrer i Cancho, R., editors, *The Evolution of Language: Proceedings of the 7th International Conference on the Evolution of Language*, pages 427–428, Barcelona. World Scientific Press.

Hoefler, S., and Smith, A. D. M. (in press). The pre-linguistic basis of grammaticalisation: A unified approach to metaphor and renalysis. To appear in *Studies in Language*.

Hoefler, S. (2008). Pragmatic plasticity: A pivotal design feature? In Smith, A. D. M., Smith, K., and Ferrer i Cancho, R., editors, *The Evolution of Language: Proceedings of the 7th International Conference on the Evolution of Language*, pages 427–428, Barcelona. World Scientific Press.

PRAGMATIC PLASTICITY: A PIVOTAL DESIGN FEATURE?

STEFAN HOEFLER
Language Evolution and Computation Research Unit, Linguistics and English Language
The University of Edinburgh, 40 George Square, Edinburgh EH8 9LL, Scotland, UK
stefan@ling.ed.ac.uk

Models developed to study the origins of language—both theoretical and computational—often tacitly assume that linguistic signals fully specify the meanings they communicate. They imply that ignoring the fact that this is not the case in actual language use is a justified simplification which can be made without significant consequences. By making this simplification, however, we miss out on the extensive explanatory potential of an empirically attested property of language: its pragmatic plasticity. In this short paper, I argue that pragmatic plasticity plays a substantial role in the evolution of language and discuss some of the key contributions this "design feature of language" (Hockett, 1960) has made to the success of linguistic communication.

Language exhibits pragmatic plasticity when the meaning a signal comes to communicate in a specific context differs from its conventional meaning—when the signal's conventional meaning under- and/or overspecifies the actually communicated meaning. Pragmatic plasticity may not be a feature pertaining to human language only, but I claim that, due to their highly developed ability to recognise common ground (Clark, 1996), it is employed by humans to a degree which cannot be found in animal communication. The same holds for conventionalisation, the process by means of which the meaning constructed in a specific context on the basis of a signal's pragmatic plasticity becomes enshrined as a new linguistic convention. The following aspects and consequences of pragmatic plasticity and its conventionalisation are thus particularly significant to language evolution:

1. *Creativity.* In effect, pragmatic plasticity is creative language use. It constitutes the major source of linguistic innovation. Theoretically, the presence of pragmatic plasticity is sufficient for language to be able to meet new communicative needs. Resorting to invention is not necessary.

2. *Adaptability.* Through pragmatic plasticity, linguistic conventions are adapted to novel contexts. This allows language to function as a communication system in the fast-changing dynamic environment of human societies. Frequently needed usages become more readily accessible—and language thus more efficient—through their conventionalisation.

3. *Expressivity.* Pragmatic plasticity means that novel meanings are expressed by using extant conventions in an under- and/or overspecified way. Once these novel usages become conventions themselves, they can exhibit pragmatic plasticity too, and thus make available yet another set of meanings not accessible before. This "ratchet-effect" (Tomasello, 1999) allows for the cumulative exploitation of ever new meaning spaces, and thus leads to a gradual increase of the number of meanings that can be expressed.

4. *Compression.* Articulation constitutes a bottleneck for linguistic communication (Levinson, 1995): meanings are transmitted via relatively slow physical channels (speech or gestures). Pragmatic plasticity accommodates this constraint by facilitating so-called lossy data compression: only information which cannot be inferred from context needs to be encoded in the linguistic signal—the rest can be left underspecified. Because we reason faster than we articulate, this increases the efficiency of linguistic communication.

5. *Symbolism.* A signal exhibits pragmatic plasticity even if it is not conventionally associated with a meaning (yet) and merely triggers the inference of meaning from the context. The conventionalisation of such maximally underspecified usage can lead to the emergence of symbolic associations.

6. *Grammaticalisation.* Pragmatic plasticity and conventionalisation are the origin of the semantic change found in grammaticalisation (Traugott & Dasher, 2005), the set of processes involved in the emergence of grammar.

7. *Ambiguity.* As it seems to be dysfunctional, ambiguity is often considered to pose an evolutionary puzzle (Hoefler, 2006). But only if we allow for ambiguity, novel usages can become conventionalised. Ambiguity is thus a crucial prerequisite for pragmatic plasticity to unfold its potential.

I conclude from these considerations that pragmatic plasticity and conventionalisation are pivotal to the emergence and evolution of language. They should therefore occupy a more central position in evolutionary linguists' models.

References

Clark, H. (1996). *Using language.* Cambridge: Cambridge University Press.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*, 88–96.

Hoefler, S. (2006). Why has ambiguous syntax emerged? In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The evolution of language* (pp. 123–130). World Scientific.

Levinson, S. C. (1995). Three levels of meaning. In F. Palmer (Ed.), *Grammar and meaning* (pp. 90–115). Cambridge: Cambridge University Press.

Tomasello, M. (1999). *The cultural origins of human cognition.* Cambridge, MA: Harvard University Press.

Traugott, E. C., & Dasher, R. B. (2005). *Regularity in semantic change.* Cambridge: Cambridge University Press.

Hoefler, S., and Smith, A. D. M. (2008). Reanalysis vs. metaphor? What grammaticalisation *can* tell us about language evolution. In Smith, A. D. M., Smith, K., and Ferrer i Cancho, R., editors, *The Evolution of Language: Proceedings of the 7th International Conference on the Evolution of Language*, pages 427–428, Barcelona. World Scientific Press.

REANALYSIS VS. METAPHOR? WHAT GRAMMATICALISATION *CAN* TELL US ABOUT LANGUAGE EVOLUTION

STEFAN HOEFLER & ANDREW D. M. SMITH

Language Evolution and Computation Research Unit, Linguistics and English Language, University of Edinburgh, 40 George Square, Edinburgh, EH8 9LL

stefan@ling.ed.ac.uk / andrew@ling.ed.ac.uk

We argue that studying grammaticalisation is useful to evolutionary linguists, if we abstract away from linguistic description to the underlying cognitive mechanisms. We set out a unified approach to grammaticalisation that allows us to identify these mechanisms, and argue that they could indeed be sufficient for the initial emergence of linguistic signal-meaning associations.

1. Introduction

Language evolution has a notorious data problem: its object of study is simply too far remote in the pre-historic past for any direct observation to be possible. In such situations, Ockham's razor recommends the assumption of the uniformity of process: that the mechanisms operating in the past are the ones still operating in the present. This would lead to the assumption that we should be able to learn something about the evolution of language from the study of language change, and in particular of semantic change leading to grammaticalisation (Heine & Kuteva, 2002a; Hurford, 2003). Grammaticalisation denotes the (unidirectional) process by which a discourse strategy, syntactic construction, or word, loses some of its independence of use and becomes more functional. It is usually accompanied by phonetic reduction and semantic bleaching and generalisation.

There is disagreement over whether the study of grammaticalisation can give useful insights into language evolution. Newmeyer (2006), for instance, criticises the assumption that the unidirectionality of grammaticalisation provides sufficient evidence that early human language contained only nouns and verbs. We argue that grammaticalisation is indeed worthy of evolutionary linguists' study, *if* one abstracts away from linguistic descriptions of individual phenomena to underlying psychological mechanisms. We thus support calls for a more cognition-oriented study of grammaticalisation (Heine, 1997; Kuteva, 2001; Tomasello, 2003):

> Exactly how grammaticalization and syntacticization happen in the concrete interactions of individual human beings and groups of human beings, and how these processes might relate to the other pro-

cesses of sociogenesis by means of which human social interaction ratchets up the complexity of cultural artefacts, requires more psychologically based linguistic research into processes of linguistic communication and language change. (Tomasello, 2003, p. 103)

The remainder of this paper falls into three sections. We first provide a unified approach to grammaticalisation, allowing us to identify the underlying cognitive mechanisms. We project these to study the emergence of a non-linguistic code, before exploring the implications of our approach for evolutionary linguistics.

2. Metaphor vs. reanalysis

Two competing kinds of accounts of grammaticalisation phenomena can be identified in the literature: those which emphasise metaphorical use (Heine, 1997), and those which emphasise reanalysis (Hopper & Traugott, 2003). We propose a unified approach based on an ostensive-inferential model of communication (Sperber & Wilson, 1995). Such a model emphasises the fact that in a given situation, a speaker and hearer assume common ground (Clark, 1996). Common ground includes, among other shared knowledge, the awareness of shared linguistic conventions and the recognition of what is relevant in a given situation, which allows the hearer to infer what the speaker intends to communicate on the basis of an ostensive stimulus provided by the speaker.

The grammaticalisation of the English construction *be going to*, which originally stood for SPATIAL MOTION, and then came to express INTENTION, as shown in Example 1, is one of the most cited examples in the grammaticalisation literature (Heine, Claudi, & Hünnemeyer, 1991; Kuteva, 2001; Hopper & Traugott, 2003), and is also a particular instance of grammaticalisation which is very common, both historically and cross-linguistically (Heine & Kuteva, 2002b).

(1) a. We are going to Windsor to see the King. (MOTION)

b. We are going to get married in June. (INTENTION, *not* MOTION)

(examples from Bybee (2003, p.147)).

We illustrate our approach by presenting the underlying psychological mechanisms, for both speaker and hearer, of metaphor- and reanalysis-based accounts of this change. In the metaphor-based scenario, detailed in Example 2, a speaker intends to express INTENTION (2a). She uses the form for SPATIAL MOTION metaphorically[a], assuming that the hearer will realise that (i) spatial motion is irrelevant in the current context, and (ii) spatial motion often implies intention, which in turn *is* relevant (2b–f). The hearer realises that the literal meaning of the

[a]There are many reasons for ad-hoc metaphorical use; these could be sociolinguistic (e.g. for prestige), or the speaker could simply not have a convention for the intended meaning in her code.

signal is irrelevant in the current context, and falls back on INTENTION, which he associates—and knows the speaker associates—with SPATIAL MOTION (2g–m).

(2) Detail of the metaphor-based scenario.

Speaker:

(a) I want to express INTENTION.
(b) I have a construction which expresses SPATIAL MOTION, and the hearer shares this convention.
(c) SPATIAL MOTION is associated with INTENTION.
(d) SPATIAL MOTION is not relevant in the given context.
(e) Because we share common ground, the hearer will be aware of (b)–(d), and realise that I am aware of it too.
(f) Because of (e), I can use the construction for SPATIAL MOTION metaphorically to convey INTENTION.

Hearer:

(g) The speaker has expressed SPATIAL MOTION.
(h) SPATIAL MOTION is not relevant in the given context.
(i) SPATIAL MOTION often implies INTENTION.
(j) INTENTION would be relevant in the given context.
(k) I must assume that the speaker is co-operative.
(l) I must also assume that the speaker is aware that I know (g)–(k), and that I know of his being aware of it.
(m) From (g)–(l), I conclude that the speaker intends to convey INTENTION.

Both speaker and hearer remember that *be going to* has been used successfully to express INTENTION; the more frequently *be going to* is used in this sense, the more deeply this new association will become entrenched (Langacker, 1987) in their knowledge. Such entrenchment eventually leads to the phenomenon of context-absorption, where a pragmatically inferred meaning becomes part of the lexical item's conventional, semantic meaning (Croft, 2000; Levinson, 2000; Kuteva, 2001; Traugott & Dasher, 2005). The entrenched meaning no longer needs to be inferred from its relevance in the given context, but can be retrieved instead from the shared conventions which make up part of language users' encyclopaedic knowledge.

In the reanalysis-based scenario, detailed in Example 3, the speaker uses *be going to* in its conventional sense to express SPATIAL MOTION—the expression of which she deems relevant in the given context (3a–e) The hearer, however, perceives things differently; he does not think that SPATIAL MOTION is relevant in the present situation but *does* believe that information about INTENTION would be

(3f). From the hearer's perspective, this appears to be exactly the same scenario as the metaphor-based scenario in Example 2. This time, the interlocutors make different adjustments to their codes: the speaker will further entrench the convention that maps *be going to* onto SPATIAL MOTION, whereas the hearer establishes a new, additional association between *be going to* and INTENTION.

(3) Detail of the renalysis-based scenario.

Speaker:

(a) I want to express SPATIAL MOTION.
(b) I have a construction for the expression of SPATIAL MOTION in my linguistic code, and the hearer shares this convention.
(c) SPATIAL MOTION is relevant in the given context.
(d) Because we share common ground, the hearer will be aware of (b)–(c) and realise that I am aware of it too.
(e) Because of (d), I can use the construction to communicate SPATIAL MOTION.

Hearer:

(f) performs the same reasoning as in (2g)–(2m) above.

A special case of the reanalysis-based scenario is one where the hearer, in the role of a language learner, does not have any existing mapping for *be going to* in his linguistic code. However, because he can work out from the context that the speaker intends to express INTENTION, he will create an association between that meaning and *be going to*. In contrast to the previous two scenarios, layering (the co-existence of an old and a new mapping, which yields polysemy) does not arise in the hearer's linguistic code in this case.

Two important conclusions can be drawn from our analysis of the metaphor- and reanalysis-based explanations of the grammaticalisation of *be going to*. First, both scenarios are based on the *same cognitive processes*: (i) those involved in ostensive-inferential communication—in particular the assumption of common ground, including knowledge of shared linguistic conventions and the recognition of what is relevant in the given context; (ii) the automatisation-based process of entrenchment. Second, the difference between the scenarios is not that only one of them uses metaphor, but rather that the (infelicitiously named) metaphor-based scenario relies on common ground having been successfully established between speaker and hearer, whereas the reanalysis-based scenario describes a situation where, although common ground is assumed by the interlocutors, there is actually a *mismatch* between their respective discourse contexts (Kuteva, 2001). The metaphor-based scenario is thus speaker-oriented, focusing on the speaker as the source of linguistic innovation, while the reanalysis-based account is hearer-oriented. Depending on which of the two perspectives one takes, however, either scenario can be regarded as a special case of the other.

3. Reconstructible meanings

How can we project these scenarios to language evolution? First, we step back to see how ostensive-inferential communication works—independent of language. We note that communication is inherently task-oriented; humans do not communicate "just so," but to do something, to achieve a goal or solve a task (Austin, 1962). The task-orientedness of communication entails that once a speaker has made manifest her intention to communicate, the hearer will have certain expectations as to what are plausible things to communicate in the given situation. In this way, a hearer discerns what is relevant from what is irrelevant in a given situation (as in the scenarios for the grammaticalisation of *going to* above), and the speaker can likewise anticipate what the hearer is likely to infer.

In the simplest case, in Fig. 1(a), making manifest one's communicative intention may suffice for the hearer to be able to infer the information one wants to communicate. The hearer's reasoning may go as follows: my conspecific exhibits behaviour that does not make sense unless she intends to communicate; therefore she intends to communicate something; in the current situation, the only thing that would make sense for her to communicate is that there is some danger around; therefore, she is communicating that there is some danger around. Note that the speaker's and hearer's assumptions can be different (i.e. there can be a contextual mismatch): if the perlocutionary effect does not differ, this may go unnoticed, and speaker and hearer will map the produced stimulus onto different utterance meanings. In Fig 1(b), for example, as long as the hearer runs and hides, it does not matter that the speaker thought she was communicating the presence of lion, while the hearer assumed that hyena were around.

Of course it is not always possible to reduce the set of plausible utterance meanings to a single one; in such cases, the hearer needs some assistance in selecting the right one, namely a *clue*. The hearer's reasoning might run along the following lines (see Fig. 1(c)). Because it does not make sense otherwise, I must interpret the speaker's behaviour as an attempt to communicate. In this situation, the only things that would make sense for her to communicate are to tell me that there is danger and to specify whether this danger is a lion or an eagle. She is communicating, so there is danger, but how can I decide if it is a lion or an eagle? The speaker must realise my dilemma, and so her ostensive stimulus will contain a clue. She is growling: lions growl, eagles don't (hyenas growl too, but this is irrelevant as there are no hyenas around at this time of year); therefore, she is communicating that there is a lion.

The cognitive mechanisms underlying these instances of communication are identical to those described in section 2 for grammaticalisation. This equivalence also extends to the entrenchment of the signal-meaning association and thus to the emergence of a convention. In all cases, the meanings which come to be associated with signals are those which can be reconstructed from the stimuli in context.

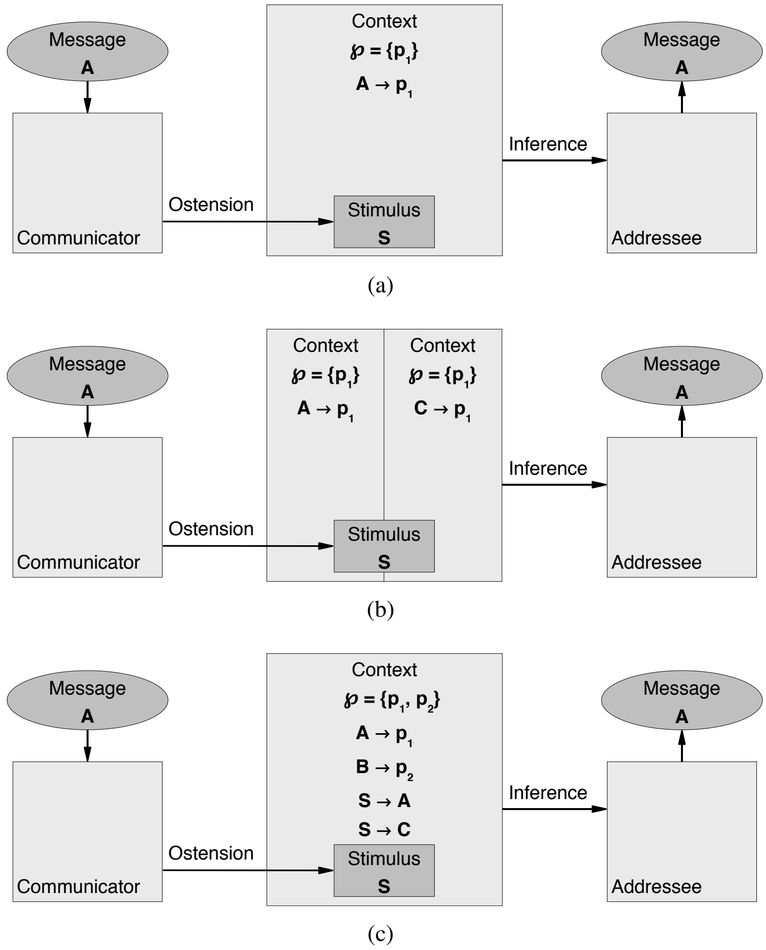


Figure 1. The reconstruction of meaning in ostensive-inferential communication, where $\wp$ is the set of plausible intended perlocutionary effects in a given situation. (a) If only one thing makes sense to be communicated, e.g. that there is some danger around (A), then the recognition of a conspecific's intention to communicate suffices to infer what she attempts to convey. (b) Contextual mismatch: the speaker means A (e.g. that there is a lion), the hearer infers C (e.g. that there is a hyena). Because both have the same perlocutionary effect p_1 (e.g. climbing a tree), the hearer's misinterpretation goes unnoticed and communication does not fail. (c) In situations where more than one thing is plausible, the speaker must additionally provide a clue. For instance, it might make sense to communicate that there is a lion (A) or an eagle (B): if there is a lion, one must climb (p_1); if there is an eagle, one must hide (p_2). Growling (S) serves as a clue: it is the sound made by lions ($S \rightarrow A$)—and by hyenas ($S \rightarrow C$), but this is irrelevant in the given context.

Every speaker innovation can only be propagated through hearer reconstruction; semantic reconstructibility therefore constrains the types of form-meaning mappings which can persist over time (Smith, 2008).

3.1. *Burling's scenario revisited*

Burling (2000) makes a case for a scenario of the emergence of linguistic symbols that is reminiscent of the reanalysis-based explanation of the grammaticalisation *be going to* we have given above. He suggests that symbols arise from situations in which one individual erroneously interprets a conspecific's behaviour as an ostensive stimulus. In our model, this would be represented as an extreme, but nevertheless ordinary, case of contextual mismatch: the hearer interprets the interaction as communicative but the speaker does not. Because the supposed ostensive stimulus will not have the properties of a proper clue, the hearer will only be able to infer a plausible meaning if there is only one relevant thing that would make sense to be communicated in the given context, and if their reaction does not expose the misunderstanding. Burling concludes that comprehension runs ahead of production: "[C]ommunication does not begin when someone makes a sign, but when someone interprets another's behaviour as a sign" (Burling, 2000, p.30). This interpretation must be rejected on the basis of our analysis of the psychological underpinnings of the equivalent reanalysis-based scenario of grammaticalisation in section 2. Although in Burling's scenario, the hearer does indeed infer something not implied by the speaker, he does so not on a whim, but under the assumption that the speaker is inviting him to make those very inferences. Rather than one being prior to the other, therefore, production and comprehension mirror each other: whatever a hearer can infer, a speaker can imply. Communication is inherently co-operative (Grice, 1975; Clark, 1996; Tomasello, 2003), and while Burling's "reanalysis-based" account cannot be ruled out, its "metaphor-based" counterpart is equally possible. Both should be seen as instances of the same set of underlying cognitive mechanisms: ostensive-inferential communication and entrenchment.

4. Conclusion

We have shown that grammaticalisation can indeed answer questions relevant to evolutionary linguists, if one moves away from linguistic classification to investigating its underlying psychological mechanisms. We have argued that the same cognitive processes that lead to grammaticalisation phenomena could also have been sufficient for the initial emergence of linguistic signal-meaning associations.

We thus neither endorse nor attempt to disprove Newmeyer (2006)'s specific criticism of the use of grammaticalisation as a source of information about language evolution. Our approach is different from both his approach and the approaches of those he criticises. We claim that the merit of studying grammaticalisation, and in fact any semantic change (Traugott & Dasher, 2005), for insights into language evolution, lies in the underlying cognitive processes it makes visible, which *can* be applied to investigate the origins of language.

References

Austin, J. L. (1962). *How to do things with words.* Oxford: Oxford University Press.

Burling, R. (2000). Comprehension, production and conventionalisation in the origins of language. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form* (pp. 27–39). Cambridge: Cambridge University Press.

Bybee, J. L. (2003). Cognitive processes in grammaticalization. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (Vol. 2). Erlbaum.

Clark, H. H. (1996). *Using language.* Cambridge: Cambridge University Press.

Croft, W. (2000). *Explaining language change: An evolutionary approach.* Longman.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts.* New York: Academic Press.

Heine, B. (1997). *Cognitive foundations of grammar.* New York: Oxford University Press.

Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework.* Chicago: University of Chicago Press.

Heine, B., & Kuteva, T. (2002a). On the evolution of grammatical forms. In A. Wray (Ed.), *The transition to language* (pp. 376–397). Oxford: Oxford University Press.

Heine, B., & Kuteva, T. (2002b). *World lexicon of grammaticalization.* Cambridge: Cambridge University Press.

Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization* (Second ed.). Cambridge: Cambridge University Press.

Hurford, J. R. (2003). The language mosaic and its evolution. In M. H. Christiansen & S. Kirby (Eds.), *Language evolution* (pp. 38–57). Oxford: Oxford University Press.

Kuteva, T. (2001). *Auxiliation: An enquiry into the nature of grammaticalization.* Oxford: Oxford University Press.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford, CA: Stanford University Press.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicatures.* Cambridge: Cambridge University Press.

Newmeyer, F. J. (2006). What can grammaticalization tell us about the origins of language? In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The Evolution of Language* (pp. 434–435). World Scientific.

Smith, A. D. M. (2008). Protolanguage reconstructed. *Interaction Studies*, *9*.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (Second ed.). Oxford: Blackwell.

Tomasello, M. (2003). On the different origins of symbols and grammar. In M. H. Christiansen & S. Kirby (Eds.), *Language evolution* (pp. 94–110). Oxford: Oxford University Press.

Traugott, E. C., & Dasher, R. B. (2005). *Regularity in semantic change.* Cambridge: Cambridge University Press.

Hoefler, S., and Smith, A. D. M. (in press). The pre-linguistic basis of grammaticalisation: A unified approach to metaphor and renalysis. To appear in *Studies in Language*.

The pre-linguistic basis of grammaticalisation: A unified approach to metaphor and reanalysis

Stefan H. Hoefler & Andrew D. M. Smith
stefan@ling.ed.ac.uk / andrew@ling.ed.ac.uk

Stefan H. Hoefler & Andrew D. M. Smith
Linguistics and English Language
The University of Edinburgh
Dugald Stewart Building, Charles Street
Edinburgh EH8 9LL, Scotland, UK

Emphasis is indicated with *italics*.

Abstract

Traditionally, grammaticalisation has been described as being based on phenomena specific to language such as metaphorical extension or reanalysis. This characterisation is somewhat in contrast to claims that grammaticalisation is involved in the much more general process of the initial emergence of language. In this article, we provide a unified analysis of both the metaphor-based and the reanalysis-based account of grammaticalisation which is grounded in the cognitive mechanisms underlying ostensive-inferential communication. We are thus able to show that the process of grammaticalisation is an instantiation of a domain-general pre-linguistic phenomenon.

1 Introduction

The way grammaticalisation has been dealt with in historical linguistics and in evolutionary linguistics is somewhat paradoxical. In historical linguistics, it is usually described as a highly specific linguistic process (Givón 1979; Bybee et al. 1994; Haspelmath 1998; Hopper & Traugott 2003), whereas in evolutionary linguistics, it is frequently invoked as an explanation for the emergence of language in general (Heine & Kuteva 2002a, 2007; Hurford 2003; Tallerman 2007). But an account of the origins of language must necessarily transcend its explanandum and provide explanations that are based on mechanisms more basic than language.

Grammaticalisation denotes the (unidirectional) process by which a discourse strategy, syntactic construction, word or morpheme loses some of its independence of use and becomes more functional in its meaning. It can be characterised as a process of semantic bleaching and generalisation accompanied by phonological reduction. Two major types of explanations have been suggested for this process in the literature: one identifies metaphorical extension as the origin of grammaticalisation (Heine et al. 1991), the other one the phenomenon of reanalysis (Hopper & Traugott 2003). Common to both approaches, however, is that they describe grammaticalisation as a relatively high-level, language-specific process. The claim that grammaticalisation constitutes one of the sources of the emergence of language in the first place (Heine & Kuteva 2002a; Hurford 2003) therefore presents somewhat of a contrast to the way it has been described in historical linguistics. Recently, it has thus even been questioned whether the process of grammaticalisation can really live up to this claim (Newmeyer 2006; see Heine & Kuteva 2007: 49–53 for a response).

In this article, we provide a unified analysis of both the metaphor-based and the reanalysis-based accounts of grammaticalisation; this unified analysis is grounded in the general cognitive mechanisms involved in ostensive-inferential communication. We are thus able to show that the process of grammaticalisation constitutes—contrary to how it has traditionally been characterised in historical linguistics—a domain-general pre-linguistic phenomenon. Our account is thereby in line with the recent shift of focus to the cognitive underpinnings of language change (for an overview of this line of research, see e.g. Evans & Green 2006) and calls for a more cognition-oriented study of grammaticalisation (Heine 1997; Kuteva 2001; Tomasello 2003) :

> "Exactly how grammaticalization and syntacticization happen in the concrete interactions of individual human beings and groups of human beings, and how these processes might relate to the other processes of sociogenesis by means of which human social interaction ratchets up the complexity of cultural artefacts, requires more psychologically based linguistic research into processes of linguistic communication and language change." (Tomasello 2003: 103)

By providing a cognition-oriented analysis of grammaticalisation, this article also relates to a number of more recent studies that explore the links between language change and language evolution (see e.g. Eckardt et al. 2008; Cooper & Kempson 2008).

The remainder of this paper falls into two main parts. We first identify the cognitive underpinnings of both the metaphor-based and the renalysis-based approaches to grammaticalisation. In the second part of the article, we then project these cognitive mechanisms to instances of pre-linguistic ostensive-inferential communication. We conclude with a brief discussion of the implications of our approach to both the study of grammaticalisation and language evolution research.

2 A unified approach

In the following discussion of the cognitive foundations of the metaphor-based and the reanalysis-based account of grammaticalisation, we use the example of the English construction *be going to* to illustrate our analysis. The grammaticalisation of *be going to* is one of the most cited examples in the grammaticalisation literature (Heine et al. 1991; Kuteva 2001; Hopper & Traugott 2003; Evans & Green 2006), and is also a particular instance of grammaticalisation which is very common, both historically and cross-linguistically (Heine & Kuteva 2002b). Originally, *be going to* stood for SPATIAL MOTION but later it came to express INTENTION and FUTURITY, as shown in example (1).

(1) a. We are going to Windsor to see the King. (MOTION)
    b. We are going to get married in June. (INTENTION/FUTURITY, *not* MOTION)

(examples from Bybee 2003: 147).

However, before we can move on to discuss the accounts of the grammaticalisation of *be going to* proposed by the metaphor-based and the reanalysis-based approach respectively, it is necessary that we provide a brief introduction of the notion of ostensive-inferential communication and its cognitive underpinnings, on which our analysis will be based.

2.1 Ostensive-inferential communication

Ostensive-inferential communication builds on the awareness of common ground. Common ground is knowledge which two interlocutors recognise as being shared in a given situation[1]. It implies (i) that the interlocutors recognise the said knowledge as shared, (ii) that they are aware that the other interlocutor recognises it as shared too, and (iii) that they realise that the other interlocutor also knows that they are aware of this (Lewis 1969; Clark 1996; Sperber & Wilson 1995). On the basis of common ground, communication can be established through an *ostensive* act performed by the communicator (a modification of the physical environment which constitutes an enhancement of the interlocutors' common ground)[2] and an *inferential* act performed by the addressee but predicted and invited by the communicator (the inference of some new information on the basis of the now altered context). The ostensive act communicates the speaker's communicative intention, by being recognised as unusual, and triggers the inferential act, in which the addressee attempts to infer the speaker's informative intention, as described by LaPolla (2006) and others.

Two special types of common ground deserve to be mentioned explicitly here. First, it is essential, for ostensive-inferential communication to be possible, that the interlocutors find a shared understanding of the assumed goal of a given interaction—which in turn builds on some understanding of each other's intentions (Tomasello et al. 2005). Such knowledge of the interactional goal enables the interlocutors to determine what is relevant in the context of that interaction; a piece of information is relevant if communicating it to the addressee would contribute to achieving the assumed goal of the interaction. Second, linguistic communication additionally makes use of yet another special type of common ground knowledge: the awareness of shared linguistic conventions. In accordance with construction-based approaches to

[1]Sperber & Wilson (1995) refer to this as mutually manifest information.

[2]An example of a basic ostensive act would be pointing at an object in order to orientate the hearer's attention to it. Diessel (1999) argues that words used exophorically, accompanying such acts, develop into demonstratives through grammaticalisation.

grammar, we understand such conventions as symbolic associations between forms and meanings (e.g. Langacker 1987; Croft 2001).

2.2 The metaphor-based scenario

Before we sketch the metaphor-based scenario of grammaticalisation, it seems appropriate to specify how we conceive the notion of 'metaphor'. What most definitions developed by metaphor theorists have in common is the core idea that metaphorical language use denotes situations where an extant linguistic form is used ("transferred" in the sense of the original Greek *metaphérein*) to express a meaning (its "metaphorical meaning") which is similar, but not identical, to the one that form is conventionally associated with (its "literal meaning") (Kövecses 2002). Metaphor is thus a form of analogy (see e.g. Hopper & Traugott (2003: 93), Fischer (2007: 122), Itkonen (2005: 41)). Lakoff & Johnson (1980) refer to the conventional meaning involved in a metaphor as its 'source', and to the the novel meaning it is used for as the 'target'. The metaphorical use of a source is possible if the semantic components that pertain only to the source but not to the intended target are irrelevant in the respective usage situation. Not all metaphors are thereby easily recognised as such. A metaphor is more conspicuous if the source shares only very few conceptual properties with the target, and less conspicuous if source and target are almost identical. Example (2) shows a very obvious case of metaphorical language use. The source (the 'chameleon') has only very few properties that also occur in the target: Sally and chameleons share their habit of frequently changing their appearance. Other features of the source, like e.g. the property of having a long tongue or the fact that chameleons are reptiles, are ignored if it is evident that they do not contribute to what the speaker plausibly wants to communicate in the given context.

(2) a. Sally is a chameleon. (*source*)
b. Sally frequently changes her appearance. (*target*)

We hold that there is no clear-cut distinction between literal use and metaphorical use but that the two rather form a continuum. Instances of literal or metaphorical use only differ in the degree of similarity between the conventional meaning of the used linguistic form and the meaning it is used to express. This view that literal use and metaphorical use—in fact any figurative language use—constitute a continuum, is advocated, for example, by Langacker (1987: 69f.), Sperber & Wilson (1995: 231–37), Carston (1997), Croft (2000: 99–114) and Wilson & Carston (2007). Metaphorical use is thus not only the basis of what we perceive as poetic language but also occurs in the small-scale deviances from convention that are ubiquitous in ordinary every-day talk (Deutscher 2005: 117ff.). It denotes any case of language use where parts of the conventional meaning of the expressed linguistic form are ignored by the interlocutors because they are not relevant in the given context. The more obviously these dismissed parts of the conventional meaning would conflict with the actually communicated meaning, the more figurative appears the respective instance of language use to be. Therefore, when we talk about metaphorical use, we refer not to any arbitrary degree of deviance from convention but to the underlying cognitive mechanism defining continuum of "loose talk" (Sperber & Wilson 1995) or "partial sanction" (Langacker 1987) that extends between literalness and poetic metaphor.

Likewise, we speak of individual instances of metaphorical extension and not of the metaphor-complexes (also called "conceptual metaphors") that e.g. Lakoff & Johnson (1980) focus on. Such extended metaphor-complexes may emerge as the cumulative effect of the repeated creation of individual metaphorical extensions that allude to already established metaphors. This statement also contains a last distinction we want to clarify: that between the *ad hoc* creation of a metaphor and so-called *dead* metaphors present in a language (Deutscher 2005). When we speak of metaphorical use, we refer to the former, namely the cognitive process by means of which a speaker uses an extant linguistic convention in a novel, metaphorical way in a particular usage event, and not to metaphors that are already established in a language and have themselves become conventional.

The reasons why speakers use extant conventions metaphorically are manifold. They may want to attract attention, establish prestige, avoid committing themselves the way they would if they used a literal expression (Pinker et al. 2008), or a metaphor may simply be shorter than a cumbersome literal circumlocution. A crucial aspect of metaphorical language use, however, is its creative potential: a metaphor may allow a speaker to express a meaning for which no extant convention exists yet. This was obviously not the case for *be going to*—English contained ways of expressing intention already before the grammaticalisation of *be going to*. But this aspect of metaphorical use will play a vital role in the examples discussed later in this article.

Along these lines, example (3) illustrates what we will call the metaphor-based scenario of grammaticalisation. A speaker intends to express INTENTION (3a). She uses the form for SPATIAL MOTION metaphorically, assuming that the hearer will realise that (i) spatial motion is irrelevant in the current context, and (ii) spatial motion often implies intention which in turn *is* relevant (3b–f)[3] The hearer realises that the literal meaning of the signal is irrelevant in the current context, and falls back on INTENTION, which he associates—and knows the speaker associates—with SPATIAL MOTION (3g–m).

(3) Detail of the metaphor-based scenario.

Speaker:

(a) I want to express INTENTION.

(b) I have a construction which expresses SPATIAL MOTION, and the hearer shares this convention.

(c) SPATIAL MOTION is associated with INTENTION.

(d) SPATIAL MOTION is not relevant in the given context.

(e) Because we share common ground, the hearer will be aware of (b)–(d), and realise that I am aware of it too.

(f) Because of (e), I can use the construction for SPATIAL MOTION metaphorically to convey INTENTION.

[Speaker expresses SPATIAL MOTION]

Hearer:

(g) The speaker has expressed SPATIAL MOTION.

(h) SPATIAL MOTION is not relevant in the given context.

(i) SPATIAL MOTION often implies INTENTION.

(j) INTENTION would be relevant in the given context.

(k) I must assume that the speaker is co-operative.

(l) I must also assume that the speaker is aware that I know (f)–(j), and that I know of his being aware of it.

(m) From (f)–(k), I conclude that the speaker intends to convey INTENTION.

[3]Of course, this implication is one of many which could be used; whether or not it is used depends on what is relevant in the current context.

2.3 Conventionalisation

The scenario described above illustrates how the original linguistic convention associated with *be going to* could have been used in a specific situation to communicate the meaning which *be going to* would later become to stand for. However, such an instance of ostensive-inferential communication is only the first step on the way to grammaticalisation. A second step needs to follow it for the process of grammaticalisation to be complete: the conveyed utterance meaning needs to become a new linguistic convention itself. This additional cognitive process of conventionalisation is initiated when the speaker and the hearer memorise the particular usage of an expression which they have just experienced. In the case described in (3), they may remember that *be going to* was used to convey intention, while at the same time maintaining their knowledge that it conventionally expresses spatial motion. The memorisation of the usage of an expression therefore has two effects: the entrenchment of a new form-meaning association, and the establishment of new common ground between the involved individuals.

The more frequently a particular expression like *be going to* is used to convey the same meaning, the more deeply the association between that form and the meaning will become entrenched in the knowledge of the user. Such entrenchment is a form of automatisation where an experienced association between a form and a meaning turns into a psychological unit and that unit becomes enforced (Langacker 1987) in the user's encyclopedic knowledge. The more deeply a form-meaning association is entrenched, the more readily it is accessible: if, for instance, the association between *be going to* and intention is sufficiently entrenched, it may be activated without the complex reasoning that was necessary to invoke it when that entrenchement had not yet happened. Through entrenchment, a form-meaning association can thus gain a certain degree of independence from its context of use. The depth of entrenchment is related to frequency of usage, because with each usage event, the association between the used linguistic form and the communicated meaning is further entrenched in both the speaker and the hearer, that is, in all individuals involved in the communicative episode (Croft 2000: 73).

While entrenchment affects each user individually, the memorisation of usages also has a more social component: it adds to the common ground that two individuals share. The awareness that *be going* *to* has in previous contexts been used successfully to convey intention, and that the hearer shares this knowledge, will allow the speaker to re-use *be going to* in the same sense in later communicative situations with the same hearer even when the context would not allow for the reasoning detailed in (3). Like entrenchment, the establishment of new common ground thus lets an association between a linguistic form and the meaning that form has been used to communicate become more context-independent. Once it has become common ground, an association between a form and a meaning can itself serve as background knowledge on the basis of which novel utterance meanings can be conveyed in future acts of ostensive-inferential communication.

The simple cognitive mechanism of usage memorisation can thus lead to the conventionalisation of a form-meaning association because it results in entrenchment and the establishment of new common ground. Ultimately, conventionalisation describes the phenomenon sometimes referred to as context-absorption (Kuteva 2001): an utterance meaning turns into a conventional meaning, that is, a pragmatically inferred meaning becomes the semantically encoded meaning associated with the used expression. The process of context-absorption—or rather its results—has been referred to frequently in both the grammaticalisation and the historical pragmatics literature (see e.g. Levinson 2000: 262–264; Kuteva 2001: 150f.; Traugott & Dasher 2005: 35).

Note that usage memorisation can bring about a situation of *layering* in an individual's linguistic knowledge, if the employed linguistic form is newly associated with a meaning that is different from its conventional meaning. The memorisation of the usage event described in (3) above, for instance, will lead to the speaker and the hearer entertaining two linguistic conventions involving *be going to*: one that associates it with SPATIAL MOTION and another one that associates it with INTENTION. However, the old and the new convention might exhibit different degrees of entrenchment. If they are both sufficiently entrenched, the impression of polysemy arises, and the expression is conceived as having multiple meanings.

2.4 The reanalysis-based scenario

As above, before we present the reanalysis-based scenario of grammaticalisation, we first explain our conception of reanalysis. Whenever a hearer interprets an utterance, he creates a set of mappings between the form and the meaning which he assumes the speaker intends to convey with her utterance. Such mappings can be of arbitrary complexity, including not only a mapping between the whole form and the whole meaning, but also mappings between individual components of the form and their semantic counterparts, as interpreted in the context of use. Every set of form-meaning mappings is thus an *analysis* of the way in which the form represents the meaning associated with it. Each independent interpretation of a form therefore yields its own analysis; if two such analyses differ in any way, then we can say that a *re-analysis* of the relationship between form and meaning has taken place.

Our notion of reanalysis is similar to, but broader than, that proposed by Hopper & Traugott (2003), who suggest that reanalysis occurs whenever "the hearer understands a form to have a structure and a meaning that are different from those of the speaker" (Hopper & Traugott 2003: 50). We concur with this, but note that reanalysis is not just restricted to comparisons between a speaker's analysis and a hearer's analysis; the comparisons can be either diachronic or synchronic, and the analyses themselves can be either by different individuals or by the same individual at different times. This view of reanalysis as a *comparison* is firmly based in the inference of meaning from context, and encompasses any difference in analysis, whether it be a difference in the meanings associated with the utterance as whole, in whether particular components of the meaning are pragmatically inferred or semantically encoded (Traugott & Dasher 2005: 35), or in how the components of the form are mapped to components of the meaning (Croft 2001: 21), including both sides of the traditional division between morphological and syntactic change (see Trask 2000, Harris & Campbell 1995, Haspelmath 1998 and McDaniels 2003 for different, more restrictive definitions of the kinds of linguistic change which count as reanalysis).

Reanalysis can also be seen as the *process* through which the two different analyses arise, and we agree with Detges & Waltereit (2002)'s argument that this process is an inevitable consequence of the cognitive mechanisms underlying ostensive-inferential communication. The context of a particular situation can be modelled, in principle, to arbitrarily complex levels, and therefore the contexts of any two situations will almost inevitably be different. In this respect, therefore, reanalysis is almost ubiquitous in communicative discourse: every time an utterance is interpreted, there will be some (possibly small) difference between the speaker's context and the hearer's context, which may lead to a mismatch between the meaning the speaker intended to communicate and that which is inferred by the hearer (Kuteva 2001: ch. 6), and thereby to differences in how the connections between components of form and components of meaning are understood.

It is important to note that because the mappings between form and meaning are internal to every linguistic individual, both they, and therefore also reanalysis itself, can never be directly observed. Differences between two sets of mappings can therefore only be uncovered through actualisation (Trask 2000), when utterances are produced which are consistent with only one of the analyses. An actualisation of the analysis of *be going to* as a marker of intention or futurity but *not* motion is shown in example (4): the utterance in 4(a) can be happily interpreted as either INTENTION or MOTION, with no change in the communicative effect, while the utterance in 4(b) cannot be intrepreted as MOTION, because this clashes with the meaning of *stay.* This clash forces the INTENTION interpretation as the only one which is communicatively plausible, and this thereby makes clear, or actualises, the analysis which maps *be going to* to INTENTION.

(4) a. I am going to play football this evening.
    b. I am going to stay here this evening.

Given this notion of reanalysis, therefore, we now set out the details of our reanalysis-based scenario in example (5), in which the speaker uses *be going to* in its conventional sense to express SPATIAL MOTION—the expression of which she deems relevant in the given context (5a–e) The hearer, however, perceives things differently; he does not think that SPATIAL MOTION is relevant in the present situation but *does* believe that information about INTENTION would be (5f-l). From the hearer's perspective, this appears to be exactly the same scenario as the metaphor-based scenario in example (3). This time, the interlocutors make different adjustments to their codes: the speaker will further entrench the convention that maps *be going to* onto SPATIAL MOTION, whereas the hearer establishes a new, additional association between *be going to* and INTENTION[4].

[4]Note that the emerging asymmetry in the interlocutors' codes can easily go unnoticed and does not necessarily lead to miscommunication. Evans & Wilkins (2000: 549f.), for instance, point out that semantic change typically occurs in what they call "bridging contexts," where "speech participants do not detect any problem of different assignments of

(5) Detail of the renalysis-based scenario[5].

Speaker:

(a) I want to express SPATIAL MOTION.

(b) I have a construction for the expression of SPATIAL MOTION in my linguistic code, and the hearer shares this convention.

(c) SPATIAL MOTION is relevant in the given context.

(d) Because we share common ground, the hearer will be aware of (b)–(c) and realise that I am aware of it too.

(e) Because of (d), I can use the construction to communicate SPATIAL MOTION.

[Speaker expresses SPATIAL MOTION]

Hearer:

(f) The speaker has expressed SPATIAL MOTION.

(g) SPATIAL MOTION is not relevant in the given context[6].

(h) SPATIAL MOTION often implies INTENTION.

(i) INTENTION would be relevant in the given context.

(j) I must assume that the speaker is co-operative.

(k) I must also assume that the speaker is aware that I know (f)–(j), and that I know of his being aware of it.

(l) From (f)–(k), I conclude that the speaker intends to convey INTENTION.

A special case of the reanalysis-based scenario is one where the hearer, in the role of a language learner, has no existing mapping for *be going to* in his lexicon. However, because he can work out from the context that the speaker intends to express INTENTION, he will create an association between that meaning and *be going to*. In contrast to the previous two scenarios, layering does not arise in the hearer's linguistic knowledge in this case.

meaning to the forms because both speaker and addressee interpretations of the utterance in context are functionally equivalent, even if the relative contributions of lexical content and pragmatic enrichment differ" (Evans & Wilkins 2000: 550).

[5]See Kuteva (2001: ch. 5–6) for a similar description of the reanalysis-based scenario.

[6]Again, this analysis is somewhat simplified, as relevance is a continuum, not a binary-valued function. It might be more accurate to state that SPATIAL MOTION is not the most relevant interpretation possible in the given context.

2.5 Comparing the two scenarios

We conclude from the analyses presented above that the two scenarios of grammaticalisation, which we have called metaphor- and reanalysis-based respectively, do not invoke different *processes* but merely describe different *circumstances*. In both scenarios, exactly the same cognitive mechanisms are at work: (i) the psychological underpinnings of ostensive-inferential communication, namely the assumption of common ground, including the knowledge of shared linguistic conventions and the recognition of what is relevant in the given context; (ii) the memorisation of usages and their subsequent entrenchment, which can lead to conventionalisation because it also involves the establishment of new common ground between the individuals participating in the communicative act.

Probably the most striking observation that arises from the comparison of the two scenarios as we have analysed them is that the hearer's part is exactly the same in the metaphor-based and in the reanalysis-based account. In both cases, the hearer assumes that the speaker has used *be going to* in a metaphorical sense. What distinguishes the two scenarios in our example is the speaker's reasoning. In the metaphor-based scenario, the speaker intends to express intention and considers it to be common ground that information about spatial motion is irrelevant in the current situation. In the reanalysis-based scenario, the speaker intends to express spatial motion—which she deems relevant in the given context. However, while in the metaphor-based scenario, speaker and hearer entertain the same assumptions about what constitutes common ground, in the reanalysis-based scenario, the hearer's assumptions on common ground are different from the speaker's. The interlocutors therefore have different beliefs (albeit ever so slightly so) about the content of the interchange, yet communication does not conspicuously fail. This is because the success of a communicative episode is determined not by comparing the inaccessible internal meanings understood by the interlocutors, but by checking whether the resulting perlocutionary effect satisfies the speaker's expectations; as long as the meaning reconstructed by the hearer has the same effect as the meaning which was intended, the difference in assumptions will pass unnoticed. To summarise, common ground is successfully established in the metaphor-based scenario, while a mismatch between the respective discourse contexts of the interlocutors occurs in the reanalysis-based scenario (for the latter, see also Kuteva

2001: ch. 6).

However, this difference between the two scenarios is not caused by different underlying cognitive processes but merely by different circumstances. Both scenarios are based on the interlocutors' capacities to entertain assumptions about what constitutes common ground in a given situation: it is only that their assumptions happen to differ in the reanalysis-based scenario. And even these two cases have to be seen as points on a continuum, namely one that ranges from situations where the interlocutors entertain identical assumptions about their common ground to situations where their assumptions are completely different. These, of course, are hypothetical limiting cases: the former condition is unlikely to occur (Sperber 1996: 101), and the latter would prevent ostensive-inferential communication from being established.

Who, then, is the innovator? In the metaphor-based scenario, both speaker and hearer end up with a novel form-meaning association: the speaker because she has used an extant convention in an innovative way, and the hearer because he has picked up the speaker's innovation. In the reanalysis-based scenario, only the hearer extends his linguistic conventions. The speaker does not use language innovatively but, because of an existing contextual mismatch, the hearer misinterprets her utterance as the innovative use of an extant convention. The two scenarios have thus merely different foci with respect to the source of innovation: one is speaker-oriented, the other one is hearer-oriented. Each of them can be seen as a special, more constrained case of the other. The reanalysis-based scenario can be interpreted as a speaker-oriented scenario where the speaker and the hearer happen to make different assumptions about what constitutes common ground and the hearer thus interprets the communicative event as innovative, when it was not intended to be. The metaphor-based scenario, on the other hand, can be understood as a hearer-oriented scenario where the interlocutors' assumptions of common ground happen to coincide and both speaker and hearer interpret the usage event they attend to as innovative. We can therefore draw the following conclusions. Language use can be either conventional or innovative—both in production and in comprehension. The scenarios of grammaticalisation we have discussed here merely represent different combinations of conventional and innovative use in the interlocutors: one and the same innovation can occur in the speaker, or in the hearer, or in both interlocutors, or in neither. This finding is summarised in Table 1. In order to complete the picture, we have also included in the table the two conditions which do not contribute to grammaticalisation: when a speaker uses language in its conventional sense and the hearer interprets it likewise (shown in the first row of the table), and when a speaker uses language innovatively but the hearer interprets it as an instance of conventional use (shown in the last row of the table). The latter could, of course, initiate grammaticalisation in the speaker's grammar, but for this innovation to spread, it would later have to occur under one of the other three sets of conditions.

[Insert Table 1 here.]

Table 1 represents a single communicative episode (e_1), of course, and in both our grammaticalisation scenarios the hearer's internalised innovation is not immediately visible in linguistic usage. Imagine a later episode (e_2) in which the hearer from e_1 acts as the speaker, and uses the innovation from the first episode; it is only during the second episode that the innovation becomes observable. The usage in the second episode is now a conventional usage from the speaker's viewpoint (who was the hearer in e_1), but innovative to the new hearer, and different from that used by the original speaker from e_1. We note that this process is equivalent to the two-stage actualisation of reanalysis, when a new analysis becomes observable through analogy Hopper & Traugott (2003: 69, 93).

Finally, we consider a condition which combines aspects of both the metaphor-based and the reanalysis-based scenarios: the speaker and hearer are both innovative in their usage, yet there is also a contextual mismatch, because they make *different* innovations, and thus different adjustments to their internal linguistic representations. In (6), the speaker intends to express INTENTION, and so uses the form for SPATIAL MOTION metaphorically, assuming that the hearer will realise that spatial motion is irrelevant, but that it often implies intention, which is relevant (6a-f). The hearer recognises that SPATIAL MOTION is irrelevant, but assumes that the form is being used metaphorically to express FUTURITY, not INTENTION. This scenario clearly illustrates the fundamentally approximate and uncertain nature of ostensive-inferential communication (Hurford 2007: 21); neither interlocutor ever knows the exact details of the other's knowledge, but must form, and act upon, assumptions made on the basis of their shared experiences.

(6) Detail of a scenario combining both metaphor-based and reanalysis-based aspects.

Speaker:

(a) I want to express INTENTION.

(b) I have a construction which expresses SPATIAL MOTION, and the hearer shares this convention.

(c) SPATIAL MOTION is associated with INTENTION.

(d) SPATIAL MOTION is not relevant in the given context.

(e) Because we share common ground, the hearer will be aware of (b)–(d), and realise that I am aware of it too.

(f) Because of (e), I can use the construction for SPATIAL MOTION metaphorically to convey INTENTION.

[Speaker expresses SPATIAL MOTION]

Hearer:

(g) The speaker has expressed SPATIAL MOTION.

(h) SPATIAL MOTION is not relevant in the given context.

(i) SPATIAL MOTION often implies FUTURITY.

(j) FUTURITY would be relevant in the given context.

(k) I must assume that the speaker is co-operative.

(l) I must also assume that the speaker is aware that I know (g)–(k), and that I know of his being aware of it.

(m) From (g)–(l), I conclude that the speaker intends to convey FUTURITY.

2.6 Multi-step scenarios

So far, we have done two things: we have sketched *possible* metaphor- and reanalysis-based pathways for the grammaticalisation of *be going to*, and we have shown that in both cases, the same cognitive mechanisms are at work. The two scenarios we have discussed illustrate under what circumstances SPATIAL MOTION could have come to express INTENTION in one step. Note that we are not claiming that motion verbs must *unavoidably* change into intention verbs, but rather that the specifics of the context is important as to the future development of a construction (see also Heine (2002); Diewald (2002)).

We show now that the same development can also have happened in more than one step, with still exactly the same set of cognitive mechanisms underlying. In fact, historical evidence suggests that this was indeed the case for the specific example of the English expression *be going to*. We must assume that the development from SPATIAL MOTION to INTENTION must have included an intermediary stage where *be going to* had come to be associated with intentional (or purposive) spatial motion, thus having assumed the notion of INTENTION while at the same time retaining its aspect of SPATIAL MOTION. Kuteva (2001: 117–121), referring to Perez (1990), points out that this development was facilitated by the intrinsically purposive nature of human spatial motion. Consequently, as Bybee et al. (1994) have shown, *be going to* (in imperfective aspect) lent itself to being complemented with a purposive clause, which solidified the conventionalisation of the aspect of INTENTION together with the original meaning SPATIAL MOTION.

In examples (7–10), we demonstrate that the individual stages of the historically attested grammaticalisation chain of *be going to* can be analysed as being initiated by exactly the same set of cognitive mechanisms that we have identified above. The stages correspond to those outlined by Kuteva (2001: 117–121). The first step, detailed in example (7), describes the addition of the aspect of INTENTION to SPATIAL MOTION[7]. The second step, detailed in example (8), illustrates the further addition of PREDICTION (FUTURITY). Both these developments describe a process of increasing semantic narrowing (SPATIAL MOTION > SPATIAL MOTION & INTENTION > SPATIAL MOTION & INTENTION & PREDICTION).

(7) Detail of step I of a multi-step scenario.

Speaker: *Henry is going to town.*

Hearer:

(a) The speaker has expressed SPATIAL MOTION.

(b) SPATIAL MOTION *alone* is not relevant in the given context.

[7] We only explicate the hearer's reasoning here, which is the same in both the metaphor-based and the reanalysis-based scenario. It is understood that, depending on which scenario one assumes for each individual step, the speaker's reasoning amounts to innovative or conventional use respectively.

(c) SPATIAL MOTION often implies INTENTION.
(d) SPATIAL MOTION together with INTENTION would be relevant in the given context.
(e) I must assume that the speaker is co-operative.
(f) I must also assume that the speaker is aware that I know (a)–(e), and that I know of his being aware of it.
(g) From (a)–(f), I conclude that the speaker intends to convey SPATIAL MOTION and INTENTION.

(8) Detail of step II of a multi-step scenario.

Speaker: *I am going to win.*

Hearer:

(a) The speaker has expressed SPATIAL MOTION and INTENTION.
(b) SPATIAL MOTION and INTENTION *alone* is not relevant in the given context.
(c) INTENTION often amounts to PREDICTION.
(d) SPATIAL MOTION and INTENTION together with PREDICTION would be relevant in the given context.
(e) I must assume that the speaker is co-operative.
(f) I must also assume that the speaker is aware that I know (a)–(e), and that I know of his being aware of it.
(g) From (a)–(f), I conclude that the speaker intends to convey SPATIAL MOTION, INTENTION and PREDICTION.

In a third and fourth step, we can observe the type of semantic broadening or bleaching that is often described as characteristic of grammaticalisation. Example (9) details a situation that initiates the use of *be going to* without the aspect of SPATIAL MOTION, and example (10) details one in which *be going to* is subsequently used without the aspect of INTENTION. The conventionalisation of the latter leads to the conventional usage of *be going to* to express (predictive) future tense. This second part of the grammaticalisation chain thus includes the following steps of semantic broadening: SPATIAL MOTION & INTENTION & PREDICTION > INTENTION & PREDICTION > PREDICTION.

(9) Detail of step III of a multi-step scenario.

Speaker: *I am going to do my best to make you happy.*

Hearer:

(a) The speaker has expressed SPATIAL MOTION, INTENTION and PREDICTION.
(b) SPATIAL MOTION is not relevant in the given context.
(c) INTENTION and PREDICTION would be relevant in the given context.
(d) I must assume that the speaker is co-operative.
(e) I must also assume that the speaker is aware that I know (a)–(d), and that I know of his being aware of it.
(f) From (a)–(e), I conclude that the speaker intends to convey INTENTION and PREDICTION only.

(10) Detail of step IV of a multi-step scenario.

Speaker: *The rain is going to come.*

Hearer:

(a) The speaker has expressed INTENTION and PREDICTION.
(b) INTENTION is not relevant in the given context.
(c) PREDICTION would be relevant in the given context.
(d) I must assume that the speaker is co-operative.
(e) I must also assume that the speaker is aware that I know (a)–(d), and that I know of his being aware of it.
(f) From (a)–(e), I conclude that the speaker intends to convey PREDICTION only.

The scenarios introduced in (3) and (5) above, which did not consider the aspect of prediction, deliberately conflated the two steps of semantic narrowing and semantic broadening into one step of semantic shift (MOTION > INTENTION), whereas the scenarios here illustrate a multi-step pathway that corresponds more closely to what we know about the specific example of *be going to* from historical evidence.

The crucial point for us, however, is this. With regard to the underlying cognitive mechanisms that we need to assume, not only does it not matter whether we adopt a metaphor-based or a reanalysis-based scenario, but it is also irrelevant how many intermediate stages we postulate: in all cases it is *exactly the same set of cognitive mechanisms* which are at work, with only the specifics of the circumstances in which a certain change is assumed to have been initiated being different.

3 The pre-linguistic basis

We now turn to the question of how our unified analysis of the process of grammaticalisation relates to the origin of language. We claim that early human pre-linguistic communication must have exhibited the same structure as the linguistic scenarios we have described. In the following paragraphs, we discuss a typical approach to language evolution in light of the above analysis of the mechanisms involved in ostensive-inferential communication. The crucial point here is to notice that ostensive-inferential communication does not need to be linguistic. Even in present-day human communication, non-linguistic ostensive-inferential communication is frequently applied where language is not available, not suitable, or not needed because there is sufficient common ground between the interlocutors (LaPolla 2006). Non-linguistic cues can be sounds, gestures or other types of behaviour that will be interpreted by an observer as communicative because they would not otherwise be plausible or relevant in a given context.

3.1 Burling's scenario revisited

One of the main questions that evolutionary linguistics has to address is how certain forms of behaviour become associated with meanings in the first place. Burling (2000, 2005) makes a case for a scenario that emphasises the role of comprehension in this process. He suggests that form-meaning associations arise when an individual erroneously interprets a conspecific's behaviour as communicative. Burling concludes that comprehension runs ahead of production in the evolution of language: "communication does not begin when someone makes a sign, but when someone interprets another's behaviour as a sign" (Burling 2000: 30). Burling's scenario is compatible with our approach—it effectively constitutes an extreme case of reanalysis—but for the following reasons, we do not find his conclusion warranted.

The central question is to identify the reasoning which Burling's scenario demands on the part of the individual who mistakenly interprets another individual's behaviour as a meaningful signal. To investigate this, we provide below the details of the comprehension process of a possible instantiation of Burling's scenario. In the interaction illustrated in example (11), one individual growls, and another individual erroneously interprets ths as a warning about the presence of a lion.

(11) Burling's reanalysis-like scenario for the origin of a pre-linguistic communicative convention.

[Individual *A* expresses GROWLING.]

Individual *B*:

(a) Individual *A* has expressed GROWLING.

(b) GROWLING itself is not relevant in the given context.

(c) GROWLING often implies LION.

(d) LION would be relevant in the given context.

(e) I must assume that individual *A* is co-operative.

(f) I must also assume that individual *A* is aware that I know (a)–(e), and that I know of his being aware of it.

(g) From (a)–(f), I conclude that individual *A* intends to convey LION.

In order to interpret the first individual's behaviour as a communicative cue, the second individual needs the set of assumptions detailed in (11a–f). These assumptions are equivalent to the hearer's reasoning in the two scenarios for grammaticalisation sketched above. We have explained above that such reasoning involves the capacity to recognise common ground (even though, as in the shown case, the respective assumptions may be misguided) and, specifically, a recognition of what information would be relevant in the given situation. Both capacities are based on the ability to interpret the intentions of other individuals.

But these cognitive capacities are at the same time sufficient for an individual to be capable of not only comprehending, but also producing an ostensive cue. It thus seems hardly plausible to assume a state where individuals were capable of comprehension but not of production, since both activities require the same minimal cognitive endowment. On the contrary, although Burling's reanalysis-like scenario is of course possible, given the mechanisms of ostensive-inferential communication and the possibility for a contextual mismatch despite the assumption of common ground, the corresponding metaphor-like scenario must be considered equally likely. In example (12), we have detailed an example of what such a scenario would look like in this case. In contrast to Burling's assumptions, our analysis suggests that the capacities to produce and comprehend ostensive cues do not emerge at different moments in human evolution. They are rather both based on the same set of more basic cognitive abilities, and once these are in place, both production and comprehension are available to an individual.

(12) Detail of a metaphor-like scenario for the origin of a pre-linguistic communicative convention.

Individual *A*:

(a) I want to express LION.

(b) I can express GROWLING.

(c) GROWLING is associated with LION.

(d) GROWLING is not relevant in the given context.

(e) Because we share common ground, individual *B* will be aware of (b)–(d), and realise that I am aware of it too.

(f) Because of (e), I can use GROWLING (quasi "metaphorically") to convey LION.

[Individual *A* expresses GROWLING.]

Individual *B*:

(g) Individual *A* has expressed GROWLING.

(h) GROWLING is not relevant in the given context.

(i) GROWLING often implies LION.

(j) LION would be relevant in the given context.

(k) I must assume that the speaker is co-operative.

(l) I must also assume that individual *A* is aware that I know (g)–(k), and that I know of his being aware of it.

(m) From (g)–(l), I conclude that individual *A* intends to convey LION.

3.2 Semantic Reconstruction

To complete the parallel with the linguistic analysis above, we present a scenario containing aspects of both metaphor-like and renalysis-like scenarios, in which both interlocutors are innovative, but they make different innovations. In example (13), one individual growls in order to warn another of the presence of a lion (13a-f), but the latter erroneously interprets it as a warning about a jackal (13g-m). Both interpretations may have the same perlocutionary effect (for instance, that both individuals climb into the trees to escape), and so the mismatch in understanding remains unnoticed.

(13) Detail of a pre-linguistic scenario that combines both metaphor-like and reanalysis-like aspects.

Individual *A*:

(a) I want to express LION.

(b) I can express GROWLING.

(c) GROWLING is associated with LION and with JACKAL.

(d) GROWLING and JACKAL are not relevant in the given context.

(e) Because we share common ground, individual *B* will be aware of (b)–(d), and realise that I am aware of it too.

(f) Because of (e), I can use GROWLING (quasi "metaphorically") to convey LION.

[Individual *A* expresses GROWLING.]

Individual *B*:

(g) Individual *A* has expressed GROWLING.

(h) GROWLING is not relevant in the given context.

(i) GROWLING often implies LION or JACKAL.

(j) JACKAL would be relevant in the given context, LION not.

(k) I must assume that individual *A* is co-operative.

(l) I must also assume that individual *A* is aware that I know (g)–(k), and that I know of his being aware of it.

(m) From (g)–(l), I conclude that individual *A* intends to convey JACKAL.

In all versions of Burling's scenario, the observed form and the inferred meaning are associated, memorised and subsequently conventionalised in the individuals' encyclopaedic knowledge bases. In the scenarios characterised by mismatches in the individuals' assumptions, however, these developing conventions also differ: in (11) only the "hearer" in fact memorises the usage and assumes that it is common ground, while in (13) both individuals memorise different usages. Because these new associations have been made by no-one else, an individual can only re-use them successfully in future situations if the context provides enough clues for the meaning to be reconstructed. The inferential reconstruction of meaning is the mechanism through which form-meaning mappings persist and survive over time; those mappings whose meanings can be easily and repeatedly reconstructed will be preferentially replicated, while those which are difficult to reconstruct will quickly perish (Smith 2008).

From the analyses presented in this section, we can see that the same general cognitive reasoning processes underpin all ostensive-inferential communication, both linguistic and pre-linguistic. These same mechanisms, which we have shown to be fundamental to the process of grammaticalisation in modern languages, could therefore also have been behind the emergence of communicative conventions, both linguistic and pre-linguistic.

There remains, of course, a big gap to be bridged between the discussed pre-linguistic forms of communication and the emergence of grammatical material in human language. Even though it is not the aim of this article to account for how this "problem of linkage," as Kirby (1999: 20) refers to this type of evolutionary puzzle, can be solved,we make a contribution toward a solution nonetheless by identifying the continual element in both the state before and the state after the transition to language. We agree with Tomasello's (2003) description of human language as the product of cumulative cultural evolution. According to this account, the human capacity to recognise common ground and in particular to understand conspecifics as intentional beings has given rise to symbolism, and once symbolic conventions were available, grammar emerged as gradually as the cumulative result of grammaticalisation and syntacticisation processes as they have been described by Givón (1979), Hopper (1987) and many others since. What this article adds to Tomasello's scenario is the recognition that the emergence of pre-linguistic symbolic conventions and the process of grammaticalisation are in fact not distinct phenomena but rather constitute instances of one and the same set of cognitive mechanisms. This analysis suggests that human language has emerged from pre-linguistic communication through iterated acts of ostensive-inferential communication, which, in the course of cumulative cultural evolution, have given rise to both symbolic conventions as well as grammatical material. Our account has the advantage that it upholds the principle of uniformity of process, which is particularly important in the light of the absence of any directly observable data on language evolution: it analyses the cognitive process underlying documented cases of grammaticalisation and projects them back to explain for the emergence of symbolic conventions.

4 Conclusion

In this paper we have examined the metaphor-based and reanalysis-based approaches to grammaticalisation in detail, and have provided a unified account based not on their linguistic properties, but on the cognitive mechanisms underlying ostensive-inferential communication, such as the assumption of common ground between interlocutors, and the memorisation of communicative experiences. Moreover, we have shown that these mechanisms are not specific to language, but are instead instances of much more fundamental, domain-general cognitive properties. Having characterised grammaticalisation in terms of domain-general mechanisms which are more basic and ancient than language, we then explored the emergence of pre-linguistic conventions using the same framework, and concluded that the same cognitive processes were also at work in this case. We therefore claim that acknowledging the cognitive mechanisms underlying ostensive-inferential communication will not only allow us to shed new light on the linguistic study of grammaticalisation, but will also provide the foundation for a unified explanation of historical language change and the origin and evolution of language.

Acknowledgements

Stefan Hoefler was funded by Cogito Foundation Research Scholarship 101/04, and Andrew Smith was funded by AHRC Grant AR/112105.

References

Burling, Robbins, 2000. Comprehension, production and conventionalisation in the origins of language. *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, Chris Knight, Michael Studdert-Kennedy, & James R. Hurford, eds., 27–39. Cambridge: Cambridge University Press.

Burling, Robbins, 2005. *The Talking Ape: how language evolved*. Studies in the Evolution of Language. Oxford: Oxford University Press.

Bybee, Joan L., 2003. Cognitive processes in grammaticalization. *The New Psychology of Language: Cognitive and functional approaches to language structure*, Michael Tomasello, ed., volume 2. Mahwah, New Jersey: Erlbaum.

Bybee, Joan L., Perkins, Revere, & Pagliuca, William, 1994. *The evolution of grammar: tense, aspect and modality in the languages of the world*. University of Chicago Press.

Carston, Robyn, 1997. Enrichment and loosening: complementary processes in deriving the proposition expressed? *Linguistische Berichte* 8:103–27.

Clark, Herbert H., 1996. *Using Language*. Cambridge: Cambridge University Press.

Cooper, Robin & Kempson, Ruth (eds.) 2008. *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*. Kings College Publications.

Croft, William, 2000. *Explaining Language Change: An Evolutionary Approach*. Harlow: Longman.

Croft, William, 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Detges, Ulrich & Waltereit, Richard, 2002. Grammaticalization vs. reanalysis: a semantic-pragmatic account of functional change in grammar. *Zeitschrift für Sprachwissenschaft* 21(S):151–195.

Deutscher, Guy, 2005. *The Unfolding of Language: an evolutionary tour of mankind's greatest invention*. New York: Metropolitan Books.

Diessel, Holger, 1999. *Demonstratives: form, function and grammaticalization*. Typological Studies in Language. John Benjamins.

Diewald, Gabriele, 2002. A model for relevant types of contexts in grammaticalization. *New Reflections on Grammaticalization*, Ilse Wischer & Gabriele Diewald, eds., Typological Studies in Language, 103–120. John Benjamins.

Eckardt, Regine, Jäger, Gerhard, & Veenstra, Tonjes (eds.) 2008. *Variation, Selection, Development: Probing the Evolutionary Model of Language Change*. Berlin: Mouton de Gruyter.

Evans, Nicholas & Wilkins, David, 2000. In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76(3):546–592.

Evans, Vyvyan & Green, Melanie, 2006. *Cognitive Linguistics: An introduction*. Edinburgh: Edinburgh University Press.

Fischer, Olga, 2007. *Morphosyntactic Change: functional and formal perspectives*. Oxford University Press.

Givón, Talmy, 1979. *On understanding grammar*. New York: Academic Press.

Harris, Alice C. & Campbell, Lyle, 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.

Haspelmath, Martin, 1998. Does grammaticalization need reanalysis? *Studies in Language* 22(2):49–85.

Heine, Bernd, 1997. *Cognitive Foundations of Grammar*. New York: Oxford University Press.

Heine, Bernd, 2002. On the role of context in grammaticalization. *New Reflections on Grammaticalization*, Ilse Wischer & Gabriele Diewald, eds., 83–101. John Benjamins.

Heine, Bernd, Claudi, Ulrike, & Hünnemeyer, Friederike, 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Heine, Bernd & Kuteva, Tania, 2002a. On the evolution of grammatical forms. *The Transition to Language*, Alison Wray, ed., 376–397. Oxford: Oxford University Press.

Heine, Bernd & Kuteva, Tania, 2002b. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd & Kuteva, Tania, 2007. *The Genesis of Grammar: A Reconstruction*. Oxford: Oxford University Press.

Hopper, P. J., 1987. Emergent grammar. *Berkely Linguistics Conference (BLS)* 13:139–157.

Hopper, Paul J. & Traugott, Elizabeth C., 2003. *Grammaticalization*. Cambridge: Cambridge University Press, second edition.

Hurford, James R., 2003. The language mosaic and its evolution. *Language Evolution*, Morten H. Christiansen & Simon Kirby, eds., chapter 3, 38–57. Oxford: Oxford University Press.

Hurford, James R., 2007. *The Origins of Meaning: Language in the Light of Evolution*. Oxford: Oxford University Press.

Itkonen, Esa, 2005. *Analogy as Structure and Process: Approaches in Linguistics, Cognitive Psychology and Philosophy of Science*. John Benjamins.

Kirby, Simon, 1999. *Function, Selection and Innateness: The Emergence of Language Universals*. Oxford: Oxford University Press.

Kövecses, Zoltán, 2002. *Metaphor: a practical introduction*. Oxford: Oxford University Press.

Kuteva, Tania, 2001. *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press.

Lakoff, George & Johnson, Mark, 1980. *Metaphors We Live By*. University of Chicago Press.

Langacker, Ronald W., 1987. *Foundations of Cognitive Grammar: Theoretical prerequisites*, volume 1. Stanford, CA: Stanford University Press.

LaPolla, Randy J., 2006. On grammatical relations as constraints on referent identification. *Voice and Grammatical Relations: Festschrift for Masayoshi Shibatani (Typological Studies in Language)*, Tasaku Tsunoda & Taro Kageyama, eds., 139–151. Amsterdam & Philadelphia: John Benjamins.

Levinson, Stephen C., 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicatures*. Cambridge: Cambridge University Press.

Lewis, David K., 1969. *Convention*. Cambridge, Massachusetts: Harvard University Press.

McDaniels, Todd, 2003. What's wrong with reanalysis? *Toronto Working Papers in Linguistics* 21:81–88.

Newmeyer, Frederick J., 2006. What can grammaticalization tell us about the origins of language? *The Evolution of Language: Proceedings of the 6th International Conference of the Evolution of Language*, Angelo Cangelosi, Andrew D. M. Smith, & Kenny Smith, eds., 434–435. Singapore: World Scientific.

Perez, A., 1990. Time in motion: Grammaticalization of the *be going to* construction in English. *La Trobe University Working Papers in Linguistics* 3.

Pinker, Steven, Nowak, Martin A., & Lee, James J., 2008. The logic of indirect speech. *Proceedings of the National Academy of Sciences* 105(3):833–838.

Smith, Andrew D. M., 2008. Protolanguage reconstructed. *Interaction Studies* 9(1):98–114.

Sperber, Dan, 1996. *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwell.

Sperber, Dan & Wilson, Deirdre, 1995. *Relevance: Communication and Cognition*. Oxford: Blackwell, second edition.

Tallerman, Maggie, 2007. Did our ancestors speak a holistic protolanguage? *Lingua* 117(3):579–604.

Tomasello, Michael, 2003. On the different origins of symbols and grammar. *Language Evolution*, Morten H. Christiansen & Simon Kirby, eds., 94–110. Oxford: Oxford University Press.

Tomasello, Michael, Carpenter, Malinda, Call, Josep, Behne, Tanya, & Moll, Henrike, 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioural and Brain Sciences* 28:675–735.

Trask, R. Larry, 2000. *The Dictionary of Historical and Comparative Linguistics*. Edinburgh: Edinburgh University Press.

Traugott, Elizabeth C. & Dasher, Richard B., 2005. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Wilson, Deirdre & Carston, Robyn, 2007. A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. *Advances in Pragmatics*, Noel Burton-Roberts, ed., 230–60. Houndmills: Palgrave MacMillan.

Table 1: The possible combinations of conventional and innovative use.

Speaker's production	Hearer's comprehension	Contextual mismatch	Scenario of grammaticalisation
conventional	conventional	no	NA
innovative	innovative	no	metaphor-based
conventional	innovative	yes	reanalysis-based
innovative	conventional	yes	NA

References

Aha, D. W. (1997). Lazy learning: Special issue editorial. *Artificial Intelligence Review*, 11:7–10.

Aha, D. W., Kibler, D., and Albert, M. K. (1991). Instance-based learning algorithms. *Machine Learning*, 6:37–66.

Aitchison, J. (1991). *Language Change: Progress or Decay?* Cambridge University Press, Cambridge.

Andersen, H. (1973). Abductive and deductive change. *Language*, 49(4):765–793.

Andersen, H. (2006). Synchrony, diachrony, and evolution. In Nedergaard Thomsen, O., editor, *Competing Models of Linguistic Change: Evolution and Beyond*, pages 59–90. John Benjamins, Amsterdam.

Arbib, M. A. (2005). From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28(2):105–124.

Arbib, M. A. and Hill, J. C. (1988). Language acquisition: Schemas replace universal grammar. In Hawkins, J. A., editor, *Explaining Language Universals*. Basil Blackwell, New York.

Atran, S. (2001). The trouble with memes: Inference versus imitation in cultural creation. *Human Nature*, 12(4):351–381.

Austin, J. L. (1962). *How to Do Things with Words*. Oxford University Press, Oxford.

Avital, E. and Jablonka, E. (2000). *Animal Traditions: Behavioural Inheritance in Evolution*. Cambridge University Press, Cambridge.

Bach, K. (2006). Impliciture vs. explicature: What's the difference? Paper delivered at the *Workshop on Explicit Communication*. Granada, May 31–June 2.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice Hall, Englewood Cliffs, New Jersey.

Banzhaf, W., Christaller, T., Dittrich, P., Kim, J. T., and Ziegler, J., editors (2003). *Advances in Artificial Life: Proceedings of the 7th European Conference on Artificial*

Life (ECAL). Number 2801 in Lecture Notes in Artificial Intelligence. Springer, Heidelberg.

Bareiss, E. R., Porter, B., and Wier, C. C. (1987). PROTOS: An exemplar-based learning apprentice. In *Proceedings of the Fourth International Workshop on Machine Learning*, pages 12–23, Irvine, California. Morgan Kaufmann.

Barlow, M. (2000). Usage, blends and grammar. In Barlow and Kemmer (2000), pages 315–346.

Barlow, M. and Kemmer, S., editors (2000). *Usage-Based Models of Language*. CSLI Publications, Stanford, California.

Barnett, H. G. (1953). *Innovation: The Basis of Cultural Change*. McGraw-Hill, New York.

Barsalou, W. L. (1990). On the indistinguishability of exemplar memory and abstraction in category representation. In Srull, T. K. and Wyer, R. S., editors, *Advances in Social Cognition*, volume 3, pages 61–88. Erlbaum, Hillsdale, New Jersey.

Basalla, G. (1988). *The Evolution of Technology*. Cambridge University Press, Cambridge.

Batali, J. (1998). Computational simulations of the emergence of grammar. In Hurford, J. R., Studdert-Kennedy, M., and Knight, C., editors, *Approaches to the Evolution of Language: Social and Cognitive Bases*, pages 405–426. Cambridge University Press, Cambridge.

Batali, J. (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In Briscoe (2002b), pages 111–172.

Bergen, B. K. and Chang, N. (2005). Embodied construction grammar in simulation-based language understanding. In Östman, J.-O. and Fried, M., editors, *Construction Grammars: Cognitive Grounding and Theoretical Extensions*, pages 147–190. John Bejamins, Amsterdam.

Bergen, B. K., Chang, N., and Narayan, S. (2004). Simulated action in embodied construction grammar. In *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society*, pages 108–113, Mahwah, New Jersey. Lawrence Erlbaum.

Bickerton, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences*, 7:173–221.

Bickerton, D. (1990). *Language and Species*. University of Chicago Press, Chicago.

Bickerton, D. (1995). *Language and Human Behaviour*. University College London Press, London.

Bickerton, D. (2000). How protolanguage became language. In Knight et al. (2000), pages 264–284.

Bickerton, D. (2003). Symbol and structure: A comprehensive framework for language evolution. In Christiansen and Kirby (2003), chapter 5, pages 77–93.

Bloomfield, L. (1984/1933). *Language*. Holt, Rinehart, and Winston, New York.

Boesch, C. and Tomasello, M. (1998). Chimpanzee and human culture. *Current Anthropology*, 39:591–614.

Boyd, R. and Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago University Press, Chicago.

Boyd, R. and Richerson, P. J. (1996). Why culture is common but cultural evolution is rare. *Proceedings of the British Acadamy*, 88:77–93.

Boyer, P. (1999). Cognitive tracks of cultural inheritance: How evolved intuitive ontology governs cultural transmission. *American Anthropologist*, 100(4):876–889.

Bresnan, J. (1982). *The Mental Representations of Grammatical Relations*. MIT Press, Cambridge, Massachusetts.

Brighton, H. (2002). Compositional syntax from cultural transmission. *Artificial Life*, 8(1):25–54.

Brighton, H. (2003). *Simplicity as a Driving Force in Linguistic Evolution*. PhD thesis, The University of Edinburgh.

Brighton, H. (2005). Linguistic evolution and induction by minimum description length. In Werning, M. and Machery, E., editors, *The Compositionality of Concepts and Meanings: Applications to Linguistics, Psychology and Neuroscience*. Ontos Verlag, Frankfurt.

Brighton, H. and Kirby, S. (2001). The survival of the smallest: Stability conditions for the cultural evolution of compositional language. In Kelemen, J. and Sosík, P., editors, *Advances in Artificial Life: Proceedings of the 6th European Conference on Artificial Life (ECAL)*, number 2159 in Lecture Notes in Computer Science, pages 592–601, Heidelberg. Springer.

Brighton, H., Kirby, S., and Smith, K. (2005). Cultural selection for learnability: Three principles underlying the view that language adapts to be learnable. In Tallerman (2005), chapter 13, pages 291–309.

Briscoe, T. (2002a). Grammatical acquisition and linguistic selection. In Briscoe (2002b), pages 255–300.

Briscoe, T., editor (2002b). *Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge University Press, Cambridge.

Briscoe, T. (2005). Co-evolution of the language faculty and language(s) with decorrelated meanings. In Tallerman (2005), pages 310–333.

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In Rosch, E. and Lloyd, B. B., editors, *Cognition and Categorization*. Erlbaum, Hillsdale, New Jersey.

Burling, R. (2000). Comprehension, production and conventionalisation in the origins of language. In Knight et al. (2000), pages 27–39.

Bybee, J. L. (1985). *Morphology: A Study into the Relation between Meaning and Form*. John Benjamins, Amsterdam.

Bybee, J. L. (1998). A functionalist approach to grammar and its evolution. *Evolution of Communication*, 2:249–278.

Bybee, J. L. (2002). *Phonology and Language Use*. Cambridge University Press, Cambridge.

Bybee, J. L. and Slobin, D. I. (1982). Why small children cannot change language on their own: Suggestions from the English past tense. In Ahlqvist, A., editor, *Papers from the 5th International Conference on Historical Linguistics*, pages 29–37, Amsterdam. John Benjamins.

Bybee, J. L. and Thompson, S. (2000). Three frequency effects in syntax. *Berkely Linguistics Society*, 23:65–85.

Campbell, L. (2001). What's wrong with grammaticalization? *Language Sciences*, 23:113–161.

Cangelosi, A., Greco, A., and Harnard, S. (2002). Symbol grounding and the symbolic theft hypothesis. In Cangelosi and Parisi (2002b), pages 191–210.

Cangelosi, A. and Parisi, D. (2002a). Computer simulation: A new scientific approach to the study of language evolution. In Cangelosi and Parisi (2002b), chapter 1, pages 3–28.

Cangelosi, A. and Parisi, D., editors (2002b). *Simulating the Evolution of Language*. Springer, London.

Cangelosi, A., Smith, A. D. M., and Smith, K., editors (2006). *The Evolution of Language: Proceedings of the 6th International Conference of the Evolution of Language*, Rome. World Scientific Press.

Carstairs-McCarthy, A. (1999). *The Origins of Complex Language: An Inquiry into the Evolutionary Beginnings of Sentences, Syllables, and Truth*. Oxford University Press, Oxford.

Carstairs-McCarthy, A. (2000). The distinction between sentences and noun phrases: An impediment to language evolution? In Knight et al. (2000), pages 248–263.

Carston, R. (1997). Enrichment and loosening: Complementary processes in deriving the proposition expressed? *Linguistische Berichte*, 8:103–127.

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Blackwell, Oxford.

Carston, R. (2004). Explicature and semantics. In Davis, S. and Gillon, B., editors, *Semantics: A Reader*, pages 817–845. Oxford University Press, Oxford.

Chandler, S. (1993). Are rules and modules necessary for explaining language? *Journal of Psycholinguistic Research*, 22(6):593–606.

Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague.

Chomsky, N. (1959). A review of Skinner's "Verbal Behavior". *Language*, 35:26–58.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1972). *Language and Mind*. Harcourt Brace Javanovich, New York.

Chomsky, N. (1980). *Rules and Representations*. Blackwell, Oxford.

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, N. (1986a). *Barriers*. MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1986b). *Knowledge of Language: Its Nature, Origin and Use*. Praeger, New York.

Chomsky, N. (1988). *Language and Problems of Knowledge*. MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1991). The semantics-pragmatics interface from different points of view. In Kasher, A., editor, *The Chomskyan Turn*, pages 5–23. Blackwell, Oxford.

Chomsky, N. (1993). A minimalist program for linguistic theory. In Hale, K. and Keyser, S. J., editors, *The View from Buidling 20*. MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1995a). Language and nature. *Mind*, 104:1–62.

Chomsky, N. (1995b). *The Minimalist Program*. MIT Press, Cambridge, Massachusetts.

Chomsky, N. (2002). *On Nature and Language*. Cambridge University Press, Cambridge.

Chouinard, M. and Clark, E. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30:637–669.

Christiansen, M. H. and Devlin, J. T. (1997). Recursive inconsistencies are hard to learn: A connectionist perspective on universal word order correlations. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, pages 113–118, Mahwah, New Jersey. Lawrence Erlbaum.

Christiansen, M. H., Ellefson, M. R., and Conway, C. M. (2002). The role of sequential learning in language evolution: Computational and experimental studies. In Cangelosi and Parisi (2002b), pages 167–187.

Christiansen, M. H. and Kirby, S., editors (2003). *Language Evolution*. Oxford University Press, Oxford.

Christy, C. (1983). *Uniformitarianism in Linguistics*. John Benjamins, Amsterdam.

Clark, H. (1996). *Using Language*. Cambridge University Press, Cambridge.

Comrie, B. (1981). *Language Universals and Linguistic Typology*. Chicago University Press, Chicago.

Comrie, B. and Kuteva, T. (2005). The evolution of grammatical structures and "functional need" explanations. In Tallerman (2005), pages 185–207.

Cook, V. J. and Newson, M. (1996). *Chomsky's Universal Grammar*. Blackwell, Oxford, second edition.

Croft, W. (2000). *Explaining Language Change: An Evolutionary Approach*. Longman, Harlow.

Croft, W. (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford University Press, Oxford.

Croft, W. and Cruse, D. A. (2004). *Cognitive Linguistics*. Cambridge University Press, Cambridge.

Cziko, G. (1995). *Without Miracles: Universal Selection Theory and the Second Darwinian Revolution*. MIT Press, Cambridge, Massachusetts.

Daelmans (1999). Memory-based language processing. *Journal for Experimental and Theoretical Artificial Intelligence*, 11(3):287–467.

Dawkins, R. (1979). *The Selfish Gene*. Oxford University Press, Oxford.

Dawkins, R. (1999). *The Extended Phenotype: The Long Reach of the Gene*. Oxford University Press, Oxford.

De Boer, B. (2002). Evolving sound systems. In Briscoe (2002b), pages 79–98.

De Boer, B. (2005). Infant-directed speech and evolution of language. In Tallerman (2005), pages 100–121.

De Pauw, G. (2003). An approximation of DOP through memory-based learning. In Bod, R., Scha, R., and Sima'an, K., editors, *Data-Oriented Parsing*. CSLI, Standford, California.

Deacon, T. W. (1997). *The Symbolic Species: The Co-Evolution of Language and the Brain*. Norton, New York.

Dennett, D. (1995). *Darwin's Dangerous Idea*. Simon and Schuster, New York.

Detges, U. and Waltereit, R. (2002). Grammaticalization vs. reanalysis: A semantic-pragmatic account of functional change in grammar. *Zeitschrift für Sprachwissenschaft*, 21:151–95.

Drachman, G. (1978). Child language and language change: A conjecture and some refutations. In Fisiak, J., editor, *Recent Developments in Historical Phonology*, pages 123–144. Mouton, Berlin.

Dressler, W. (1974). Diachronic puzzles for natural phonology. In Bruck, A., Fox, R. A., and LaGaly, M. W., editors, *Papers from the Parasession on Natural Phonology*, pages 95–102, Chicago. Chicago Linguistic Society.

DuBois, J. (1987). The discourse of ergativity. *Language*, 68:805–855.

Dunbar, R. I. M. (1998). *Grooming, Gossip and the Evolution of Language*. Harvard University Press, Cambridge, Massachusetts.

Eddy, J. C. (2005). Iterated learning: The exemplar-based learning approach. Master's thesis, The University of Edinburgh, Edinburgh.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*. MIT Press, Cambridge, Massachusetts.

Evans, V. and Green, M. (2006). *Cognitive Linguistics: An Introduction*. Edinburgh University Press, Edinburgh.

Fillmore, C. J. (1975). An alternative to checklist theories of meaning. In *Proceedings of the First Annual Meeting of the Berkeley Linguistic Society*, pages 123–131, Amsterdam. North Holland.

Fillmore, C. J., Kay, P., and O'Connor, M. K. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language*, 64:501–538.

Futuyma, D. J. (1998). *Evolutionary Biology*. Sinauer Associates, Sunderland, Massachusetts, third edition.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, 29:737–767.

Galef, B. G. (1988). Imitation in animals: History, definition, and interpretations of data from the psychological laboratory. In Zentall, T. and Galef, B. G., editors, *Social Learning: Psychological and Biological Perspectives*, pages 3–29. Lawrence Erlbaum, Hillsdale, New Jersey.

Galef, B. G. (1992). The question of animal culture. *Human Nature*, 3:157–178.

Galef, B. G. (1996). Social enhancement of food preferences in Norway rats: A brief review. In Hayes, C. M. and Galef, B. G., editors, *Social Learning in Animals*, pages 49–64, San Diego. Academic Press.

Garrod, S., Fay, N., Lee, J., Oberlander, J., and MacLeod, T. (2007). Foundations of representation: Where might graphical symbols systems come from? *Cognitive Science*, 31(6):961–987.

Gazdar, G., Klein, E., Pullum, G., and Sag, I. (1985). *Generalized Phrase Structure Grammar*. Blackwell, Oxford.

Gell-Mann, M. (1992). Complexity and complex adaptive systems. In Hawkins, J. A. and Gell-Mann, M., editors, *The Evolution of Human Languages*. Addison-Wesley, Reading, Massachusetts.

Gil, D. (2006). Early human language was isolating-monocategorial-associational. In Cangelosi et al. (2006), pages 91–98.

Givón, T. (1971). Historical syntax and synchronic morphology: An archaeologist's field trip. *Chicago Linguistic Society*, 7:394–415.

Givón, T. (1979). *On Understanding Grammar*. Academic Press, New York.

Givón, T. (1990). *Syntax: A Functional-Typological Introduction*. John Benjamins, Amsterdam.

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

Goldin-Meadow, S. (2008). Gesture, speech, and language. In Smith et al. (2008), pages 427–428.

Gould, S. J. (1997). Darwinian fundamentalism. *New York Review of Books*, pages 34–52.

Greenberg, J. H. (1963). *Universals of Language*. MIT Press, Cambridge, Massachusetts.

Grice, H. P. (1957). Meaning. *Philosophical Review*, 66:377–388.

Grice, H. P. (1968). Utterer's meaning, sentence-meaning, and word-meaning. *Foundations of Language*, 4:225–242.

Grice, H. P. (1975). Logic and conversation. In Cole, P. and Morgan, J., editors, *Syntax and Semantics 3: Speech Acts*. Academic Press, New York.

Haiman, J. (1980a). Dictionaries and encyclopedias. *Lingua*, 50:781–819.

Haiman, J. (1980b). The iconicity of grammar: Isomorphism and motivation. *Language*, 56:515–540.

Haiman, J. (1983). Iconic and economic motivation. *Language*, 59:781–819.

Haiman, J. (1985). *Natural Syntax*. Cambridge University Press, Cambridge.

Haiman, J. (1994). Ritualization and the development of language. In Pagliuca (1994), pages 3–28.

Haiman, J. (2008). In defence of iconicity. *Cognitive Linguistics*, 19(1):35–48.

Hall, C. (1992). *Morphology and Mind: A Unified Approach to Explanation in Linguistics*. Routledge, London.

Halle, M. (1962). Phonology in generative grammar. *Word*, 18:54–72.

Harris, A. C. and Campbell, L. (1995). *Historical Syntax in Cross-Linguistic Perspective*. Cambridge University Press, Cambridge.

Hashimoto, T. (2002). The constructive approach to the dynamic view of language. In Cangelosi and Parisi (2002b), pages 307–324.

Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language*, 22(2):315–351.

Hauser, M., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298:1569–1579.

Hawkins, J. A. (1994). *A Performance Theory of Order and Constituency*. Cambridge University Press, Cambridge.

Hawkins, J. A. (2004). *Efficiency and Complexity in Grammars*. Oxford University Press, Oxford.

Heine, B. (1997). *Cognitive Foundations of Grammar*. Oxford University Press, New York.

Heine, B., Claudi, U., and Hünnemeyer, F. (1991). *Grammaticalization: A Conceptual Framework*. University of Chicago Press, Chicago.

Heine, B. and Kuteva, T. (2002). *World Lexicon of Grammaticalization*. Cambridge University Press, Cambridge.

Henrich, J., Boyd, R., and Richerson, P. J. (in press). Five misunderstandings about cultural evolution. In Sperber, D., editor, *Epidemiology of Ideas*. Open Court.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, 93:411–428.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203.

Hoefler, S. (2006a). Modelling pragmatic aspects of language evolution: The case of the articulation bottleneck. Paper presented at the *Annual Meeting of the Linguistics Association of Great Britain (LAGB 2006)*, Newcastle upon Tyne.

Hoefler, S. (2006b). Why has ambiguous syntax emerged? In Cangelosi et al. (2006), pages 123–130.

Hoefler, S. (2008). Pragmatic plasticity: A pivotal design feature? In Smith et al. (2008), pages 427–428.

Hoefler, S. and Smith, A. D. M. (2008). Reanalysis vs. metaphor? What grammaticalisation *can* tell us about language evolution. In Smith et al. (2008), pages 427–428.

Hoefler, S. and Smith, A. D. M. (in press). The pre-linguistic basis of grammaticalisation: A unified approach to metaphor and reanalysis. *Studies in Language*.

Holland, J. H. (2005). Language acquisition as a complex adaptive system. In Minett, J. W. and Wang, W. S.-Y., editors, *Language Acquisition, Change and Emergence*, pages 411–436. City University of Hong Kong Press, Hong Kong.

Hooper, J. B. (1980). Child morphology and morphophonemic change. In Fisiak, J., editor, *Historical Morphology*, pages 157–187. Mouton, Berlin.

Hopper, P. J. (1987). Emergent grammar. *Berkeley Linguistics Conference (BLS)*, 13:139–157.

Hopper, P. J. (1991). Functional explanations in linguistics and the origins of language. *Language and Communication*, 1–2:45–47.

Hopper, P. J. and Traugott, E. C. (2003). *Grammaticalization*. Cambridge University Press, Cambridge, second edition.

Hopper, R., Knapp, M. L., and Scott, L. (1981). Couples' personal idioms: Exploring intimate talk. *Journal of Communication*, 31:23–33.

Horn, L. and Ward, G., editors (2004). *The Handbook of Pragmatics*. Blackwell, Oxford.

Horn, L. R. (2004). Implicature. In Horn and Ward (2004), pages 3–28.

Huang, Y. (2007). *Pragmatics*. Oxford University Press, Oxford.

Hull, D. L. (1988). *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago, Chicago University Press.

Hurford, J. R. (1987). *Language and Number: The Emergence of a Cognitive System*. Blackwell, Oxford.

Hurford, J. R. (1990). Nativist and functional explanations in language acquisition. In Roca, I. M., editor, *Logical Issues in Language Acquisistion*, pages 85–136. Foris, Dordrecht.

Hurford, J. R. (2000). Social transmission favours linguistic generalisation. In Knight et al. (2000), chapter 19, pages 324–352.

Hurford, J. R. (2002). Expression/induction models of language evolution. In Briscoe (2002b), chapter 10.

Hurford, J. R. (2003). The language mosaic and its evolution. In Christiansen and Kirby (2003), chapter 3, pages 38–57.

Hurford, J. R. (2004). Human uniqueness, learned symbols and recursive thought. *European Review*, 12(4):551–565.

Hymes, D. (1972). Competence and performance in linguistic theory. In Huxley, R. and Ingram, E., editors, *Language Acquisition: Models and Methods*. Academic Press, New York.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences*, 3(7):272–279.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press, Oxford.

Jackendoff, R. and Pinker, S. (2005). The nature of the language faculty and its implications for the evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition*, 97:211–225.

Jespersen, O. (1922). *Language, its Nature, Development and Origin*. London. Reprinted by W. W. Norton, 1996.

Kawai, M. (1965). Newly-acquired pre-cultural behavior of the natural troop of Japanese monkeys on Koshima Islet. *Primates*, 6:1–30.

Kawamura, S. (1959). The process of sub-culture propagation among Japanese macaques. *Primates*, 2:43–60.

Kay, P. and Fillmore, C. (1999). Grammatical constructions and linguistic generalizations: The *What's X doing Y* construction. *Language*, 75:1–34.

Keller, R. (1994). *On Language Change: The Invisible Hand in Language*. Routledge, London.

Keller, R. (1998). *A Theory of Linguistic Signs*. Oxford University Press, Oxford.

Kemmer, S. and Barlow, M. (2000). A usage-based conception of language. In Barlow and Kemmer (2000), pages vii–xxviii.

Kinsella, A. R. (in press). *Language Evolution and Syntactic Theory*. Cambridge University Press, Cambridge.

Kirby, S. (1997). Competing motivations and emergence: Explaining implicational hierarchies. *Language Typology*, 1:5–31.

Kirby, S. (1999). *Function, Selection and Innateness: The Emergence of Language Universals*. Oxford University Press, Oxford.

Kirby, S. (2000). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In Knight et al. (2000), chapter 18, pages 303–323.

Kirby, S. (2001). Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation*, 5(2):102–110.

Kirby, S. (2002a). Learning, bottlenecks and the evolution of recursive syntax. In Briscoe (2002b), chapter 6.

Kirby, S. (2002b). Natural language from artificial life. *Artificial Life*, 8(2):185–215.

Kirby, S., Dowman, M., and Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *PNAS*, 104(12):5241–5245.

Kirby, S. and Hurford, J. R. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In Cangelosi and Parisi (2002b), chapter 6, pages 121–148.

Kirby, S., Smith, K., and Brighton, H. (2004). From UG to universals: Linguistic adaptation through iterated learning. *Studies in Language*, 28(3):587–607.

Knight, C., Studdert-Kennedy, M., and Hurford, J. R., editors (2000). *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge University Press, Cambridge.

Koton, P. (1988). Reasoning about evidence in causal explanations. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, pages 256–261, St. Paul, Maine. Morgan Kaufmann.

Kristiansen, K. (2005). Innovation and invention: Independent event or historical process? In Renfrew and Bahn (2005), pages 151–155.

Kummer, H. and Goodall, J. (1985). Conditions of innovative behaviour in primates. *Philosophical Transactions of the Royal Society of London*, B308:203–214.

Kuteva, T. (2001). *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford University Press, Oxford.

Labov, W. (1982). Building on empirical foundations. In Lehmann, W. P. and Malkiel, Y., editors, *Perspectives on Historical Linguistics*, pages 17–92. John Benjamins, Amsterdam.

Lakoff, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press, Chicago.

Lakoff, G. (1991). Cognitive versus generative linguistics: How commitments influence results. *Language and Communication*, 1–2:53–62.

Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press, Chicago.

Langacker, R. W. (1977). Syntactic reanalysis. In Li (1977), pages 57–139.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites*, volume 1. Stanford University Press, Standford, California.

Langacker, R. W. (1991). *Foundations of Cognitive Grammar: Descriptive Application*, volume 2. Stanford University Press, Standford, California.

Langacker, R. W. (2000). A dynamic usage-based model. In Barlow and Kemmer (2000), pages 1–64.

Laurence, S. and Mangolis, E. (2001). The poverty of the stimulus argument. *British Journal of the Philosophy of Science*, 52:217–276.

Levinson, S. C. (1995). Three levels of meaning. In Palmer, F. R., editor, *Grammar and Meaning*, pages 90–115. Cambridge University Press, Cambridge.

Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicatures*. Cambridge University Press, Cambidge.

Lewis, D. K. (1969). *Convention: A Philosophical Study*. Harvard University Press, Cambridge, Massachusetts.

Li, C. N., editor (1977). *Mechanisms of Syntactic Change*. University of Texas Press, Austin.

Lightfoot, D. (1979). *Principles of Diachronic Syntax*. Cambridge University Press, Cambridge.

Lightfoot, D. (1991). *How to Set Parameters*. MIT Press, Cambridge, Massachusetts.

Livingstone, D. (2002). The evolution of dialect diversity. In Cangelosi and Parisi (2002b), pages 99–118.

Marcus, G. (1993). Negative evidence in language acquisition. *Cognition*, 46:43–85.

Maynard Smith, J. and Harper, D. (2003). *Animal Signals*. Oxford University Press, Oxford.

McCarthy, J. (1990). Formalization of two puzzles involving knowledge. In Lifschitz, V., editor, *Formalizing Common Sense: Papers by John McCarthy*, pages 158–166. Ablex Publishing, Norwood, New Jersey.

McDaniels, T. (2003). What's wrong with reanalysis? *Toronto Working Papers in Linguistics*, 21:81–88.

McMahon, A. M. S. (1994). *Understanding Language Change*. Cambridge University Press, Cambridge.

Medin, D. L. and Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85:207–238.

Mehler, J., Sebastian, N., Altmann, G., Dupoux, E., Christophe, A., and Pallier, C. (1993). Understanding compressed sentences: The role of rhythm and meaning. *Annals of the New York Academy of Sciences*, 682:272–282.

Miller, G. F. (2000). *The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature*. Doubleday, London.

Milroy, J. (1992). *Linguistic Variation and Change*. Blackwell, Oxford.

Milroy, J. (1993). On the social origins of language change. In Jones, C., editor, *Historical Linguistics: Problems and Perspectives*, pages 215–236. Longman, London.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill, New York.

Mühlhäusler, P. (1991). Comments on Newmeyer's "Functional explanation in linguistics and the origins of language". *Language and Communication*, 1–2:75–78.

Mundinger, P. (1980). Animal cultures and a general theory of cultural evolution. *Ethology and Sociobiology*, 1:183–223.

Nänny, M. and Fischer, O., editors (1999). *Iconicity in Language and Literature*. John Benjamins, Amsterdam.

Nerlich, B. and Clarke, D. D. (1992). Outline of a model for semantic change. In Kellermann, G. and Morrissey, M. D., editors, *Diachrony within Synchrony: Language History and Cognition. Papers from the Interantional Symposium at the*

University of Duisburg, 26–18 March 1990, pages 125–141, Frankfurt am Main. Peter Lang.

Newmeyer, F. J. (1991). Functional explanation in linguistics and the origin of language. *Language and Communication*, 1(2):3–28.

Newmeyer, F. J. (2003). What can the field of linguistics tell us about the origins of language? In Christiansen and Kirby (2003), chapter 4, pages 58–76.

Niyogi, P. (2002). Theories of cultural evolution and their application to language change. In Briscoe (2002b), pages 205–233.

Niyogi, P. and Berwick, R. C. (1995). The logical problem of language change. Technical report. MIT AI Memo No. 1516.

Niyogi, P. and Berwick, R. C. (1996). A language learning model for finite parameter spaces. *Cognition*, 61(1):161–193.

Niyogi, P. and Berwick, R. C. (1997). Evolutionary consequences of language learning. *Linguistics and Philosophy*, 20:697–719.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, 15:39–57.

Nunberg, G., Sag, I. A., and Wasow, T. (1994). Idioms. *Language*, 70:491–538.

O'Brien, M. and Lyman, R. L. (2000). *Applying Evolutionary Archaeology*. Plenum, New York.

O'Brien, M. and Lyman, R. L. (2002). Evolutionary archeology: Current status and future prospects. *Evolutionary Anthropology*, 11:26–36.

Odling-Smee, F. J., Laland, K. N., and Feldman, M. W. (2003). *Niche Construction: The Neglected Process in Evolution*. Princeton University Press, Princeton, New Jersey.

Origgi, G. and Sperber, D. (2000). Evolution, communication and the proper function of language. In Carruthers, P. and Chamberlain, A., editors, *Evolution and the Human Mind*, pages 140–169. Cambridge University Press, Cambridge.

Origgi, G. and Sperber, D. (2004). A pragmatic perspective on the evolution of language and languages. In Dominey, P. F. F., Origgi, G., and Reboul, A., editors, *Interdisciplines Conference: Issues in Coevolution of Language and Theories of Mind*, Lyon. Institut des Sciences Cognitives.

Pagliuca, W., editor (1994). *Perspectives on Grammaticalization*. John Benjamins, Amsterdam.

Parker, A. R. (2006). Evolving the narrow language faculty: Was recursion the pivotal step? In Cangelosi et al. (2006), pages 239–246.

Peirce, C. S. (1932). Harvard University Press, Cambridge, Massachusetts.

Piatelli-Palmarini, M. (1989). Evolution, selection, and cognition: From "learning" to parameter setting in biology and the study of language. *Cognition*, 31:1–44.

Pinker, S. (1997). *How the Mind Works*. Norton, New York.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language*. Harper Collins, New York.

Pinker, S. (2003). Language as an adaptation to the cognitive niche. In Christiansen and Kirby (2003), chapter 2, pages 16–37.

Pinker, S. and Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Science*, 13(4):707–784.

Pollard, C. and Sag, I. A. (1993). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Pullum, G. and Scholz, B. (2002). Empirical assemssment of stimulus poverty arguments. *The Linguistic Review*, 19:9–50.

Recanati, F. (1989). The pragmatics of what is said. *Mind and Language*, 4:295–329.

Recanati, F. (2004a). *Literal Meaning*. Cambridge University Press, Cambridge.

Recanati, F. (2004b). Pragmatics and semantics. In Horn and Ward (2004), pages 442–462.

Renfrew, C. and Bahn, P., editors (2005). *Archaeology: The Key Concepts*. Routledge, London.

Rissanen, J. and Ristad, E. S. (1994). Language acquisition in the MDL framework. In Ristad, E. S., editor, *Language Computation*. American Mathematical Society, Philadelphia.

Rissland, E. L., Kolodner, J., and Waltz, D. (1989). Case-based reasoning *from* DARPA: Machine learning program plan. In *Proceedings of the Case-Based Reasoning Workshop*, pages 1–13, Pensecola Beach, Florida. Morgan Kaufmann.

Roberts, M., Onnis, L., and Chater, N. (2005). Acquisition and evolution of quasi-regular languages. In Tallerman (2005), pages 334–356.

Rosenbach, A. and Jäger, G. (2008). Priming and unidirectional language change. *Theoretical Linguistics*, 34(2):85–113.

Saussure, F. d. (1959/1916). *Course in General Linguistics*. Philosophical Library, New York.

Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language*, 24:139–161.

Saxton, M., Houston-Rice, C., and Dawson, N. (2005). The prompt hypothesis: Clarification requests as corrective input for grammatical errors. *Applied Psycholinguistics*, 26:393–414.

Sayood, K. (2006). *Introduction to Data Compression*. Morgan Kaufmann (Elsevier), Amsterdam, third edition.

Schiffer, S. R. (1972). *Meaning*. Oxford University Press, Oxford.

Scholz, B. and Pullum, G. (2006). Irrational nativist exuberance. In Stainton, R. J., editor, *Contemporary Debates in Cognitive Science*, pages 59–80. Blackwell, Oxford.

Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press, Chicago.

Shennan, S. (2002). *Genes, Memes and Human History: Darwinian Archaeology and Cultural Evolution*. Thames & Hudson, London.

Shennan, S. (2005). Cultural evolution. In Renfrew and Bahn (2005), pages 49–54.

Simone, R., editor (1995). *Iconicity in Language*. John Benjamins, Amsterdam.

Skousen, R. (1989). *Analogical Modelling of Language*. Kluwer Academic, Dordrecht.

Skousen, R. (1992). *Analogy and Structure*. Kluwer Academic, Dordrecht.

Slobin, D. I. (1997). *The Crosslinguistic Study of Language Acquisition*, volume 4. Lawrence Erlbaum Associates, Mahwah, New Jersey.

Smith, A. D. M. (2003a). Intelligent meaning creation in a clumpy world helps communication. *Artificial Life*, 9:175–190.

Smith, A. D. M. (2003b). Semantic generalisation and the inference of meaning. In Banzhaf et al. (2003), pages 499–506.

Smith, A. D. M. (2005a). The inferential transmission of language. *Adaptive Behavior*, 13(4):311–324.

Smith, A. D. M. (2005b). Mutual exclusivity: Communicative success despite conceptual divergence. In Tallerman (2005), pages 372–388.

Smith, A. D. M. (2006a). Semantic reconstructibility and the complexification of language. In Cangelosi et al. (2006), pages 307–314.

Smith, A. D. M. (2008). Protolanguage reconstructed. *Interaction Studies*, 9(1):100–116.

Smith, A. D. M., Smith, K., and Ferrer i Cancho, R., editors (2008). *The Evolution of Language: Proceedings of the 7th International Conference on the Evolution of Language*, Barcelona. World Scientific Press.

Smith, E. A. (2000). Three styles in the evolutionary analysis of human behaviour. In Cronk, L., Chagnon, N., and Irons, W., editors, *Adaptation and Human Behaviour*, pages 27–46. de Gruyter, New York.

Smith, E. E. and Medin, D. L. (1981). *Categories and Concepts*. Harvard University Press, Cambridge, Massachusetts.

Smith, K. (2003c). Learning biases for the evolution of linguistic structure: An associative network model. In Banzhaf et al. (2003), pages 517–524.

Smith, K. (2004). The evolution of vocabulary. *Journal of Theoretical Biology*, 228(1):127–142.

Smith, K. (2006b). Cultural evolution of language. In Brown, K., editor, *The Encyclopedia of Language and Linguistics*, pages 315–322. Elsevier, second edition.

Smith, K. (2006c). The protolanguage debate: Bridging the gap? In Cangelosi et al. (2006), pages 315–322.

Smith, K., Brighton, H., and Kirby, S. (2003). Complex systems in language evolution: The cultural emergence of compositional structure. *Advances in Complex Systems*, 6(4):537–558.

Solan, Z., Ruppin, E., Horn, D., and Edelman, S. (2005). Evolution of language diversity: Why fitness counts. In Tallerman (2005), pages 357–372.

Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*. Blackwell, Oxford.

Sperber, D. and Wilson, D. (1995). *Relevance: Communication and Cognition*. Blackwell, Oxford, second edition.

Sperber, D. and Wilson, D. (1997). Précis of *Relevance: Communication and Cognition*. In Kahser, A., editor, *Pragmatics: Critical Concepts*, volume 5, pages 82–115. Routledge, London.

Stalnaker, R. C. (1978). Assertion. In Cole, P., editor, *Syntax and Semantics 9: Pragmatics*, pages 315–332. Academic Press, New York.

Stanfill, C. and Waltz, D. (1986). Categories and concepts. *Communications of the ACM*, 29:1213–1228.

Steels, L. (1996a). Emergent adaptive lexicons. In Maes, P., Mataric, M., Meyer, J.-A., Pollok, J., and Wilson, S., editors, *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behaviour*, Cambridge, Massachusetts. MIT Press.

Steels, L. (1996b). Perceptually grounded meaning creation. In Tokoro, M., editor, *Proceedings of the International Conference on Multi-Agent Systems*, pages 338–344, Cambridge, Massachusetts. MIT Press.

Steels, L. (1997). Constructing and sharing perceptual distinction. In van Someren, M. and Widmer, G., editors, *Proceedings of the European Conference on Machine Learning*, pages 4–13, Berlin. Springer.

Steels, L. (1999). *The Talking Heads Experiment*. Laboratorium, Antwerp.

Steels, L. (2002). Grounding symbols through evolutionary language games. In Cangelosi and Parisi (2002b), pages 211–226.

Steels, L. (2004). Constructivist development of grounded construction grammars. In Daelemans, W., editor, *Proceedings of the Annual Meeting of Association for Computational Linguistics*.

Steels, L. and Kaplan, F. (2002). Bootstrapping grounded word semantics. In Briscoe (2002b), pages 53–74.

Sterelny, K. (2006). The evolution and evolvability of culture. *Mind and Language*, 21(2):137–165.

Stokoe, W. C. (1978). *Sign Language Structure: The First Linguistic Analysis of American Sign Language*. Linstok Press, Silver Spring, Maryland.

Stokoe, W. C. (2005/1960). Sign language structure: An outline of the visual communication systems of the American deaf. *Journal of Deaf Studies and Deaf Education*, 10(1):3–37.

Tallerman, M., editor (2005). *Language Origins: Perspectives on Evolution*. Oxford University Press, Oxford.

Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua*, 117(3):579–604.

Taylor, J. R. (2002). *Cognitive Grammar*. Oxford University Press, Oxford.

Teal, T. K. and Taylor, C. E. (2000). Effects of compression on language evolution. *Artificial Life*, 6(2):129–143.

Tellier, I. (1999). Towards a semantic-based theory of language learning. In *Proceedings of the 12th Amsterdam Colloquium*, pages 217–222.

Tellier, I. (2000). Semantic-driven emergence of syntax : the pinciple of compositionality upside-down. In *Proceedings of the 3rd Conference on the Evolution of Language*, pages 220–224, Paris.

Thompson, S. A. (1991). On addressing functional explanation in linguistics. *Language and Communication*, 1(2):93–96.

Timberlake, A. (1977). Reanalysis and actualization in syntactic change. In Li (1977), pages 141–177.

Tomasello, M. (1995). Language is not an instinct. *Cognitive Development*, 10:131–156.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press, Cambridge, Massachusetts.

Tomasello, M. (2002). Some facts about primate (including human) communication and social learning. In Cangelosi and Parisi (2002b), chapter 15, pages 327–340.

Tomasello, M. (2003a). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press, Cambridge, Massachusetts.

Tomasello, M. (2003b). On the different origins of symbols and grammar. In Christiansen and Kirby (2003), chapter 6, pages 94–110.

Tomasello, M. (2004). What kind of evidence could refute the UG hypothesis? *Studies in Language*, 28(3):642–645.

Tomasello, M., Kruger, A. C., and Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16:495–552.

Trask, R. L. (2000). *The Dictionary of Historical and Comparative Linguistics*. Edinburgh University Press, Edinburgh.

Traugott, E. C. and Dasher, R. B. (2005). *Regularity in Semantic Change*. Cambridge University Press, Cambridge.

Vihman, M. M. (1980). Sound change and child language. In Traugott, E. C., Labrum, R., and Shepherd, S., editors, *Papers from the 4th International Conference on Historical Linguistics*, pages 303–320, Amsterdam. John Benjamins.

Wedgwood, D. (2007). Shared assumptions: Semantic minimalism and Relevance Theory. *Journal of Linguistics*, 43:647–681.

Weinreich, U., Labov, W., and Herzog, M. I. (1968). Empirical foundations for a theory of language change. In Lehmann, W. P. and Malkiel, Y., editors, *Directions for Historical Linguistics: A Symposium*, pages 97–195. University of Texas Press, Austin.

Wheeldon, L. R. and Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language*, 34:311–334.

Wierzbicka, A. (1972). *Semantic Primitives*. Athenäum, Frankfurt am Main.

Wierzbicka, A. (1989). Semantic primitives and lexical universals. *Quaderni di Semantica*, pages 103–121.

Wilson, D. and Carston, R. (2007). A unitary approach to lexical pragmatics. In Burton-Roberts, N., editor, *Advances in Pragmatics*, pages 230–260. Parlgrave MacMillan, Houndmills.

Wolff, J. G. (1982). Language acquisition, data compression and generalization. *Language and Communication*, 2:57–89.

Wood, M. M. (1993). *Categorial Grammar*. Routledge, London.

Wray, A. (2000). Holistic utterances in protolanguage: The link from primates to humans. In Knight et al. (2000), chapter 17, pages 285–303.

Wray, A. (2002). Conceptualizing transition in an evolving field. In Wray, A., editor, *The Transition to Language*, chapter 1, pages 1–20. Oxford University Press, Oxford.

Wray, A. (2005). "Needs only analysis" in linguistic ontogeny and phylogeny. In *Proceedings of the Second International Symposium on the Emergence and Evolution of Linguistic Communication*, pages 151–160.

Zipf, G. K. (1968/1935). *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. MIT Press, Cambridge, Massachusetts.

Zuidema, W. (2003). How the poverty of the stimulus solves the poverty of the stimulus. In Becker, S., Thrun, S., and Obermayer, K., editors, *Advances in Neural Information Processing Systems 15 (Proceedings of NIPS'02)*, Cambridge, Massachusetts. MIT Press.