Chapter 1

Introduction

1.1 Language and evolutionary theory

There is something about language and evolutionary theory that has appealed to scientists ever since the days of Darwin. In fact, Darwin himself was one of the first to notice ‘curious parallels’ between the formation and development of languages and species (Darwin 1989 [1877]: 94). Since then, the topic has been a recurring one, especially in studies on language change (e.g. Schleicher 1983 [1869], Sapir 1923, Keller 1994, Lass 1997, Croft 2000, Janda & Joseph 2003, Verhagen 2004).

What, then, are these ‘curious parallels’? First, there is the observation that languages show many parallels with biological species: they change continuously, new languages can form because of lack of contact between two parts of a population, and languages can become extinct. This means that there is something to language that makes it behave in an evolutionary fashion. Second, there are similar mechanisms at work in both language and species, such as transmission from parents to children, the presence of selection pressures favoring one variant over the other (e.g. Verhagen 2000a), adaptation and presumably exaptation (Lass 1990). This means it might be possible to use evolutionary mechanisms to explain linguistic phenomena.

The latter observation will be the starting point of this thesis, which results from a project aimed at creating closer collaboration between the fields of biology and linguistics. More specifically, the aim of this thesis is to look at concrete cases of language change and try to describe and possibly explain them by using a biological, that is, evolutionary, viewpoint. As such, it can be seen as an application of the theory of evolutionary linguistics as introduced by Croft (2000).

1.2 Recent evolutionary approaches in linguistics

Serious interest in the use of evolutionary concepts in language has increased considerably in the last two decades, in different ways. First, the transition to language – which is commonly referred to as ‘language evolution’ – has become an established field of research (Hurford, Studdert-Kennedy & Knight 1998, Hauser,
Chomsky & Fitch 2002, Wray 2002, Pinker & Jackendoff 2005). Language is considered to be a trait in evolutionary terms, and at some point in the evolution of man, early humans have acquired this trait. This fact gives rise to many important questions that are studied in this field. For example, what was so different about early humans when compared to other species in that only they acquired language? What were the first phases of language like? Why has it developed the way it has?

In the study of language evolution, computer modeling has become an important subfield (Steels 1998, Nowak & Krakauer 1999, Vogt 2005, and various papers in Briscoe 2002, Cangelosi & Parisi 2002 and Christiansen & Kirby 2003). In many of these studies, the development of language in a human population is simulated with models that are often referred to as ‘agent-based’ models. These ‘agents’ represent humans with different kinds of cognitive and social behavior that interact with each other in a population. They have some sort of linguistic knowledge, which is shaped on the basis of the input they receive in the interaction with other agents. This set-up makes it possible to simulate different scenarios of language evolution and thus get a better insight into which factors have played an important role in this process.

While most of these simulation studies focus on the transition to language, similar computer models are also used to study the dynamics of language in a fully-grown system, or, in other words, language change. Although the focus is different, both types of models are based on the same rationale: the linguistic knowledge of individuals is (partly) formed by interaction with other individuals. By simulating these interactions, it is possible to get a better understanding of the mechanisms that play a role in the dynamics of this process. Examples of these studies are Niyogi & Berwick (1997), Nettle (1999), Yang (2000) and Baxter et al. (2006, to appear).

Parallel to the increased interest for the evolution of language is the increased interest in the use of evolutionary theory outside biology. Studies such as Cavalli-Sforza & Feldman (1981) and Boyd & Richerson (1985) present mathematical models to apply evolutionary concepts to culture. These models are based on the assumption that evolution is a general phenomenon that is not restricted to biological systems. Instead, evolution is a logical outcome of any system in which there is variation and differential replication of variants, disregarding what is being replicated. This means that culture, in the broad sense of learned traits, is subject to evolution as well: cultural traits such as ideas and beliefs are passed on from individual to individual and show variation. A profound discussion of the implications of this view can be found in Hull (1988) and Dennet (1995).

This notion of a general theory of evolution takes the evolution of cultural traits like ideas, habits, beliefs and also language beyond the idea of ‘parallelism’ or ‘analogy’ with regards to biological evolution, and makes it possible to study them in a more formal way in its own right. The transmission of cultural traits occurs when one individual learns from another, not by the transmission of genetic
material. This difference in the mechanism of transmission does not alter the fact that culture in this sense constitutes a system of information that is inherited by a new generation from the previous one. Cultural traits are also subject to mutation in that copying errors might occur or that traits can be created or altered by individuals. Again, the difference in mechanisms does not alter the fact that culture is an inherited system of information that can change. The existence of both mutation and reproduction results in a system of variation and inheritance, which can lead to evolutionary change in combination with drift or by selection.

Finally, evolutionary concepts are also increasingly used in the field of language classification. This field studies the question how different languages from the same language family are related. It is based on the assumption that languages gradually change. When two languages are compared, one will therefore find differences between the two. If a whole set of languages are compared, one will find that some languages differ more from one another than others. Another assumption is that linguistic similarity is a measure for relatedness: languages that are closely related are more similar than languages that are only distantly related. This means that, by measuring their similarity, it is possible to reconstruct how languages within a language family are related. This approach is called phylogenetic reconstruction.

Of course, language classification itself is not a new enterprise in linguistics. However, a new development in the field in the last decade, is the introduction of computational techniques that were originally developed for biology (e.g. Gray & Atkinson 2003, and Dunn, Terrill, Reesink, Foley & Levinson 2005). Computers are used to reconstruct how a group of languages are related by comparing a large set of characters from each language. These characters are usually ‘basic’ words like ‘hand’ or ‘father’, but they can be structural characteristics as well. The phylogenetic tree is then calculated by comparing the number of cognates that are shared by each set of two languages.

One of the chapters of this thesis is dedicated to phylogenetic reconstruction. However, instead of using these techniques to classify languages within a larger language family, I will use them to study change within a single language. In other chapters, I will use agent-based models like the ones described earlier to study specific cases of language change. This means that I will restrict myself to the evolution of fully developed linguistic systems, and not consider the transition to language.

1.3 A framework of evolutionary linguistics

The agent-based models that I present in this thesis are in many ways similar to the ones I discussed in the previous section: they consist of a population of agents who exchange linguistic utterances, and whose individual behavior can be regulated. A
difference between them lies in the underlying theoretical basis of the models. While many of the aforementioned models use the theory of generative grammar, I will use a usage-based approach to language. In this section, I will discuss this approach, and its link with evolutionary theory, in more detail.

The usage-based approach to language is a general term for a view of language and grammar in which language and its structural characteristics are considered to be a product of usage over time (Bybee 2006). When compared to the generative approach to language, this means a move away from language as a system in isolation, with usage being peripheral, to a system in space and time, with usage as the core aspect. Speakers are assumed to be continuously sensitive to the language around them: they shape their linguistic knowledge on the basis of the input they receive in communication. This means they are also sensitive to the frequency with which they perceive linguistic input.

This usage-based approach lies at the basis of different, yet related linguistic theories on e.g. grammar (cognitive and construction grammar (Langacker 1987, Croft & Cruse 2004), radical construction grammar (Croft 2001), language acquisition (Tomasello 2003), exemplar theory (e.g. Daelemans 1998, Pierrehumbert 2001), language change (Traugott & Dasher 2002, Verhagen 2004) and grammaticalization theory (Heine, Claudi & Hünnemeyer 1991, Bybee, Perkins & Pagliuca 1994, Hopper & Traugott 2003). It is also used by Croft (2000), who presents an evolutionary framework of language and language change. His theory is based on the idea that evolution is a general phenomenon that is present in all kinds of systems with certain characteristics, as I mentioned above (e.g. Cavalli-Sforza & Feldman 1981, Hull 1988).

The core notion of Croft’s theory is that (diachronic) patterns in the language of a population can be explained as a result of the behavior of individuals in that population. This approach is very similar to the view on language change as proposed by Keller (1994). In his ‘invisible hand theory’, Keller states that language change is the unintended result of intentional individual behavior. Individuals deliberately use certain strategies when they communicate, and these strategies, or ‘maxims’, have both a communicative and social basis. Examples are ‘talk in such a way that you are communicatively successful’, or ‘talk in such a way that you do not spend superfluous energy’. In other words, individual users select certain linguistic variants over others because they want to be communicatively successful or they want to talk as economically as possible. This selection, in turn, can, but does not necessarily have to, lead to linguistic change. Language users, therefore, can change their language without intending to do so: their intentions are aimed at more ‘local’ goals such as communicative and social success.

Although Keller does not explicitly present his theory as being ‘evolutionary’, it has many similarities to evolutionary theory. First, the phenomenon of language change is reduced to actions at the level of the individual.
Second, there is variation of linguistic items and differential selection of these items through the use of the maxims. Transmission of the items takes place in communication. There is one important issue in which the invisible hand theory differs from evolutionary theory: the latter models change as a two-step process: there is the generation of variation (mutation or innovation) on the one hand, and the possible selection and spread of mutations or innovations through a population on the other hand. In Keller’s theory, however, such a strict distinction of the two processes is absent; change is modeled as a one-step process: as the cumulative effect of individual actions.

Croft’s theory of evolutionary linguistics can be seen as an elaboration of the theory of Keller. Croft argues that the *utterance* is the unit of transmission and selection, and that a language can be defined as a population of utterances in a speech community (Croft 2000: 26). Utterances are replicated, or transmitted, in communication between individuals within a speech community. By default, utterances are produced and perceived that are already part of the linguistic convention of the speech community, in which case the language remains the same. However, innovations might occur, intentionally or unintentionally, during communication, and when these innovations spread through the speech community, (parts of the) language will change.

An example of an innovation is when a speaker uses an utterance in a way that has not been used before, e.g. the use of *going to* to express intention instead of a change in location. Similarly, innovation might also occur on the side of the hearer. In the case of *going to*, an utterance in which a change in location was intended by the speaker might be understood by the hearer as mainly expressing intention.

Croft’s point that innovations can both be intentional and unintentional is interesting, because the nature of the innovations might affect the direction of change. In different chapters in this thesis, I will discuss this topic in more detail. Whether an innovation spreads through a community and leads to a proper linguistic change depends, according to Croft, solely on social factors like speaker status or population structure (Croft ibid.: 178ff). Whether or not this sharp social delimitation of spread, or propagation, is desirable could be debated. The use of an innovation also leads to differences in entrenchment in the head of the hearer, which in turn could lead to altered use. Eventually, this could lead to propagation of the innovation through the speech community without the ‘help’ of any social factors.

Croft’s theory of evolutionary linguistics, with the utterance as the unit of transmission, is a theory that puts language use in the center of the linguistic system. Transmission takes places when individuals communicate with each other, and utterances can be subject to innovation and selection. As such, the approach differs from generative approaches to language change. In a generative framework, individuals in a generation *t* have a particular linguistic knowledge in the form of a
grammar. This grammar is used in utterances in communication and as such, passed on to individuals from the next generation $t + 1$. In turn, these individuals have to construct their grammar on the basis of these utterances, but because of the critical period, their input is limited. This limitation can lead to linguistic knowledge in generation $t + 1$ that differs from that of the original generation, and if this is the case, language change has occurred. In this approach, transmission and innovation takes place in the process of language acquisition, and selection can occur through the succession of generations of speakers (Croft ibid.: 44). In the usage-based framework, change can also occur within a single generation of speakers, because individuals continuously adapt their linguistic knowledge on the basis of the utterances they perceive in communication. Therefore, in the usage-based framework, the main mechanisms of change are to be found in the factors that affect the transmission of utterances between individuals, instead of those that affect the transmission of knowledge from one generation to another.

1.4 Advantages of an evolutionary approach to language

There are several aspects of the evolutionary linguistic framework that make it appealing for historical linguistic research. As I mentioned earlier, complex phenomena at the population level are taken as the result of individual behavior and the interactions between individuals, which are easier to study and to understand. Also, instead of only looking for system-specific causes when studying linguistic phenomena, patterns and developments in language can now be studied as possibly resulting from general mechanisms as well. Not only will this increase our understanding of these processes, it also makes it possible to apply knowledge from outside linguistics to the field.

Another appealing aspect of an evolutionary framework is that it offers a coherent model of language that can deal with both language change and stasis (or: lack of change). Language is considered an inherently dynamic system that is constantly in motion, but in which this dynamicity does not necessarily have to lead to significant changes. Linguistic variants can come into existence in a range of different ways (e.g. by mutations that occur during production or in the process of perception), but whether or not these variants spread through a population depends on a range of other factors. For example, one allophone might have a selectional advantage over another in that it differs more clearly from other phonemes in the language. If this allophone spreads through the population, there is evolution through selection. Social factors might also lead to selectional advantages for a particular linguistic variant, for example if the individual that uses it has some sort of social prestige in the population. On the other hand, variants can also spread
through a population by chance. In this case, evolution occurs as a result of random drift. The evolutionary framework makes it possible to study these phenomena, and to get a better understanding of the role of the different relevant factors in the process within a single methodology.

The use of evolutionary theory in linguistic research also allows for a strongly quantitative, mathematical view on language change, which enables the use of computer modeling. This tool is increasingly used in historical linguistic research, and the simulation of language use and language change has become an important subfield of linguistics in recent decades. In this respect, linguistics has the advantage over genetics in that our historical knowledge of language exceeds our historical knowledge of genes. Thus, it is possible to use linguistic data in the construction of the models, and in the interpretation of their results.

Approaches differ with respect to the view on language, the linguistic phenomenon that is studied, and the focus on language origin or language change. However, most studies using computer simulations are similar in their basic set-up of the model, which consists of a (small) population of individuals, or agents, who produce and perceive the linguistic item under investigation. By focusing on particular aspects of either individual behavior or population structure, the development of the item in question is then investigated. In this sense, these models could be called ‘evolutionary’, because they investigate language as a product of transmission and selection in a population of agents.

Although computer simulations are a useful tool in the study of language and language evolution, it is necessary to also consider their limitations. Models are supposed to be a simplified representation of real-world phenomena, which, in this case, is language. However, this representation is always based on a number of assumptions about the system that is simulated. Furthermore, the model’s behavior cannot always be understood precisely, because of its complex nature. This does not necessarily imply that the model’s results cannot lead to a better understanding of the phenomenon that is investigated. On the contrary, its use makes it possible to study the role of different factors that cannot be studied as easily otherwise. Still, it does mean that it is necessary to relate both the model’s set-up and its results back to empirical findings as much as possible, and to always remain cautious about the behavior it shows.

1.5 This thesis

As I mentioned earlier, the general goal of this thesis is to apply an evolutionary viewpoint to a number of concrete cases of language change, and to see how such an approach can lead to a better understanding, and perhaps to an explanation, of the changes in question. I will use computer simulations to study these cases. Regarding
linguistic theory, I will use Croft’s framework of evolutionary linguistics (Croft 2000) and basic insights from the usage-based approach (Bybee 2006).

First of all, a series of choices had to be made about the topics in this thesis. The first choice was to focus on semantics. There are two reasons for this choice. First, human language is not unique as a vocal communication system; different species are known to also have a relatively developed vocal communication system that is culturally transmitted: bats, whales and dolphins (also known as ‘cetaceans’), seals, sea lions, and hummingbirds, songbirds and parrots (see Fitch 2005 for an overview). Of these species, songbirds are the largest and best-studied group (e.g. Marler & Slabbekoorn 2004). However, human language stands out because it appears to be the only communication system with an elaborate, open-ended semantic component (Nonhuman primates are known to have quite complex minds, but they ‘lack a communicative mechanism that is capable of expressing most of this mental activity’ (Fitch 2005: 206)). It is therefore interesting to study how changes in this ‘unique’ dimension can be explained in general evolutionary terms; after all, not only the forms but also the meanings are transmitted through cultural processes (that is, learning).

Second, the modeling of semantics is somewhat of a challenge, because meaning is not transmitted directly in communication. When someone perceives an utterance, its phonetic and much of the morphosyntactic structure are directly clear, but its meaning is not. This means that the semantics of an utterance has to be reconstructed on the basis of indirect information. This might be one of the reasons that most of the computer models so far have focused on syntax and phonetics, while, as Steels (2003) points out, unidirectionality and grammaticalization (changes in the function of elements towards that of a grammatical item) have received less attention in this line of research.

Another choice that had to be made was the language of the case studies, and I have chosen Dutch. Of course, choosing a ‘small’ language is a disadvantage for many readers because the material that will be discussed is much less accessible to them. At the same time, the field of historical linguistics is strongly English-centered, and although this is not a bad thing per se, I believe it is good to focus on other languages as well. Also, I believe that a study of change in a particular language demands detailed knowledge of the nuances of the language that only native speakers readily have access to.

The last choice concerned the focus when using computer models. These models allow for testing of the role of every possible factor in language change, which is both their strong and their weak point. It is very important to decide on a limited number of relevant factors to investigate beforehand, in order not to drown in a sea of information. It is therefore also important to keep the model simple; with too many factors involved, it becomes difficult to understand the role of each single
factor in the process as a whole, and this would defeat the purpose of using computer models altogether.

A commonly used distinction is that between so-called ‘functional’ and ‘social’ factors. The former affect the individual behavior of the agent, such as its production, perception, and the way it processes its linguistic experiences. Although these factors also possibly involve the hearing party (while we are dealing with language in use), they can be called ‘agent-internal’ in order to distinguish them from factors that are concerned with phenomena at a group level. These are social factors, such as population structure, age and social status. For this thesis, I have decided to focus mainly on ‘functional’ factors. Note, however, that this does not mean that I consider social factors not to be important in language change.

In the first chapter of this thesis, chapter 2, I introduce a general computer model of semantic change. With the use of this model, I discuss how certain tendencies of semantic change, such as the change from lexical to functional meaning, can be explained by very general factors such as frequency of use.

Chapter 3 deals with the principle of competitive exclusion in language. I argue that this biological principle can be used to explain linguistic competition phenomena as well, such as two forms competing for the same meaning. To illustrate this, I discuss the linguistic equilibrium situations of two variants of adjective-noun combinations, compounds and phrases, in Dutch, German and English. Using a computer model, I argue that competition does not necessarily have to lead to extinction of one of the two forms, if there is a possibility for at least partial semantic differentiation.

Chapters 4-6 all deal with the development of the Dutch verb *krijgen* (comparable to English *get*). In chapter 4, I introduce the verb with a diachronic corpus study, on the basis of which I give a detailed presentation of its syntactic and semantic development. I also discuss the mechanisms that arguably underlie these developments.

I use these findings in chapter 5, in which I present a computer model of *krijgen*. This model simulates the semantic development of the verb and combines insights from usage-based and exemplar-based theories in an evolutionary framework. I focus on the notion of preservation and loss of original meaning in verbs like *krijgen*, which show a great amount of polysemy.

In chapter 6, I introduce a new methodology for historical linguistic research. This method is called phylogenetic reconstruction and is widely used in biology in the reconstruction of historical relations between species. Recently, it has also been introduced in the classification of languages within language families like Indo-European. I now discuss the application of the method in the study of the development of a linguistic item within a single language, that of *krijgen*. The major
benefit of this method is that a historical reconstruction can be obtained on the basis of synchronic material.

I conclude this thesis with a discussion of the findings from the respective chapters.