11 Language as a complex adaptive system: the interaction of cognition, culture and use

11.1 Typology and universals

A linguistic theory must strive to be applicable to all human languages and thus must recognize at some level what all languages have in common. Generative theory, for instance, has sought commonalities in the form of universals of grammar at the level of phrase structure rules and conditions and constraints on movement rules. While there are certainly many tendencies and repeated patterns cross-linguistically, stating universals at this level has largely been unsuccessful in accounting for the empirical data (see Newmeyer 2005). In this chapter we will consider tracing the tendencies and patterns observable across languages to the interaction of the cognitive processes that have been discussed in the previous chapters of this book. This approach allows us to integrate synchronic patterns with patterns of language change and provides the framework for forming a more comprehensive theory that explains the range of structures found in the languages of the world. But in addition to accounting for similarities among languages, it is also important to account for major typological differences. Towards this goal it is suggested here, following other research (Perkins 1992, Wray and Grace 2007) that cultural factors may come into play. Indeed social and cultural factors have remained in the background in the previous discussion, but clearly such factors cannot be ignored in a full account of the emergence of language.

Construction grammar as studied by Fillmore and colleagues emphasizes the idiomaticity of grammar. The burgeoning study of prefabs and formulaic language mentioned earlier (Wray 2002, and others) also emphasizes the extent to which linguistic knowledge is specific to particular words and phrases and therefore to particular languages. Radical Construction Grammar (Croft 2001) is typological in orientation but emphasizes the specifics of particular constructions within and across languages, arguing against static universals – for instance, on the level of 'the passive construction'. Following the Greenbergian tradition, this approach considers the universals to be evident in the way constructions develop over time. The new approaches to child language acquisition that were mentioned in earlier chapters also emphasize the child's use of

specific instances of constructions at early stages, leading up gradually to more general constructions (Lieven et al. 1997, Tomasello 2003, Dbrowska and Lieven 2005). Despite the emphasis on the language-specific and local generalizations, none of these researchers would deny that all human languages are very similar; not only is this a firm intuition shared by researchers who work on language, but it is backed up by extensive demonstration of similarities across unrelated languages.

Usage-based theory developed directly out of, and is in a sense just a new name for, American functionalism, which has been practiced for many decades (Noonan 1998). The first usage-based linguist of the twentieth century was Joseph Greenberg. Although he is better known for his studies in typology and universals, he also showed an interest in frequency effects in helping to explain cross-linguistic patterns (Greenberg 1966). Other usage-based linguists who are also typologists include T. Givón (1975, 1979), Sandra Thompson (1988, 1998), Paul Hopper (Hopper and Thompson 1980, 1984), John Haiman (1985), and William Croft (2003). These researchers connect their dual interests in usage patterns and typology with the theoretical proposition that frequently occurring usage patterns will be reflected in cross-linguistically common grammatical patterns. Note that this approach, initiated by Greenberg in the 1960s, has the central properties of a complex adaptive systems approach: it postulates a small number of factors interacting locally from which emerge a dynamic, apparently global structure. Specifically, some concepts from complexity theory apply to the usage-based approach to language in the broadest sense – that is, to language as a world-wide phenomenon encompassing all known types. Thus the repeated grammaticalization paths alluded to in Chapter 6 can be thought of as 'strange attractors' in the sense that certain cycles seem to be repeated across languages and across time but without ever being precisely identical.

The goal of this chapter, then, is to present this view with several examples, showing that commonalities across languages can be explained with reference to language use as filtered through the processing mechanisms discussed in previous chapters. However, our first topic will be the general one of how cross-linguistic similarities can be incorporated into a theory of language.

11.2 Cross-linguistic similarities in a theory of language

A basic question for any linguistic theory to address is the nature of the human genetic endowment that makes language possible. Perhaps the most fundamental consideration is whether language similarities are to be attributed to domain-general or domain-specific processes and abilities. As we have noted in earlier chapters, 'domain-general' abilities are those that are

also used outside of language – in general cognition – and include chunking, categorization, the use of symbols, the ability to make inferences and so on. 'Domain-specific' abilities would be those that are specific to language and not evidenced elsewhere. Aspects of the ability to process speech auditorily, for example, may turn out to be quite specific to language and not a process used elsewhere.

Among the domain-specific abilities that have been proposed one might distinguish between structural knowledge and processing tendencies. The innate parameters of generative grammar would be structural knowledge – specific knowledge about how languages are structured. An example of a processing constraint that might be innate would be the parsing constraint discussed in Hawkins 2009. The structural knowledge would become manifest during the acquisition of language by children, and the processing constraints would affect choices of structures and thus affect grammar through usage.

In earlier chapters I have argued against structural knowledge as innate. In Chapter 6 I also argued against the view that children play a more significant role than adults in changing language. The view that structural properties of language are innate requires that these properties appear in the acquisition process and that any changes in structure occur in this process. If we find that children in fact are not the major instigators of linguistic change, then the main link between innate universals and language structure cannot be established. In addition, I have demonstrated in several chapters of this book that the categories and constituents of grammar are gradient in their properties and change as language is used. The view that such properties are innate is not compatible with this demonstrated gradience. Finally, the fact that all categories and structures evolve gradually from other categories and structures also argues against the existence of static and innate universals of language (Bybee, 2009a and 2009b).

In fact it is more parsimonious to accept the challenge to derive language from non-language – that is, from domain-general principles – and to view language as a complex adaptive system. First, if we adhere to the assumption that the processes underlying language are specific to language, we will never discover if any of them applies outside of language. In contrast, if we start by examining the possibility of domain-general processes in operation, then we can eventually distinguish the domain-general from domain-specific. Second, in a complex-systems approach 'universals' or similarities across languages are emergent and dynamic, not static and given. Such a view is more in keeping with the facts: there are very few absolute universals; rather there are many possible typological patterns at every level and few pure types. These facts are consistent with the hypothesis that there are multiple factors involved in shaping a language.

11.3 Synchronic observations, diachronic paths and domain-general processes

American Structuralists, with their highly empirical orientation, emphasized the differences among languages and took great pains not to reduce all languages to the mold of what Benjamin Lee Whorf called Standard Average European (Whorf 1941; see also Sapir 1921). In contrast, in the middle of the twentieth century, both Noam Chomsky and Joseph Greenberg began to emphasize the many commonalities shared by distinct languages. Their approaches were, however, quite different: Chomsky postulated an innate 'Universal Grammar' as a starting point, so that 'universals' could be discovered by studying only a single language (Chomsky 1965). In contrast Joseph Greenberg studied hundreds of languages to establish their similarities and differences (Greenberg 1963, 1966, 1978a, 1978b). Based as it is on an understanding of the diversity as well as similarity among languages, Greenberg's theory is much more subtle and nuanced than Chomsky's in that it sees relations among the properties of languages that can be expressed in hierarchies and implicational statements, all of which are ultimately based on diachronic changes. In addition Greenberg tries to get closer to explanations for similarities while generativists seem satisfied with the a priori postulation that there are innate linguistic universals. In a Chomskian theory universals are properties that cannot be learned from experience; in Greenberg's theory or other usage-based theories patterns that turn out to be cross-linguistically similar are indeed learned from experience.

The generative idea of universals is carried to its logical conclusion in Optimality Theory where linguistic forms are derived by the application of universal constraints. Since the constraints can override one another, none of the 'universals' is absolute in its application. This fits with the facts – there are very few absolute universals – but it also makes the theory untestable. Practitioners of this theory play fast and loose with the empirical data, postulating 'universal constraints' on the basis of very little data, often on the basis of patterns that recur in a few related languages (see Bybee 2005).

The other common property of Universal Grammar found in work within generative theory and Optimality Theory is the view that structural properties of language are universal and innately given. That means that observational generalizations about grammar – distributions of consonants and vowels or ordering of subject, verb and object – are proposed to be a part of linguistic competence without any further search for the principles that underlie these observed generalizations. This is another reason why the Greenbergian approach is more sophisticated and satisfying. In the Greenbergian approach one does not stop at the observation stage, but rather one continues to piece together a sequence of diachronic developments that lead to both the similarities and the differences among languages (Greenberg 1969, 1978a and 1978b, Croft 2001,

Bybee 2006b). The ultimate goal is to identify the causal mechanisms that lead to the paths of development that underlie the observed regularities.

For instance, Greenberg 1978b discusses a diachronic continuum among languages (even some unrelated languages) from demonstratives to noun class markers. This continuum is a grammaticalization path, as demonstratives in some languages develop into noun class markers (e.g. Bantu languages). Many such paths have been discovered in the extensive literature on grammaticalization. They are paths of semantic development (paralleled by paths of development of form) that can be found approximated in different, related or unrelated languages across time (Givón 1979, Bybee and Dahl 1989, Bybee et al. 1994, Heine et al. 1991, Heine and Kuteva 2002). Just to take one example for which the empirical data are quite clear, there are many languages with a future form that is derived from a verb or construction meaning 'movement towards a goal'. In the seventy-six-language sample of Bybee et al. 1994, in which the languages studied were maximally unrelated genetically, the following languages were found to have a future marker built from a movement verb: Margi (Chadic), Tucano (Andean-Equatorial), Guaymí (Chibchan), Danish (Indo-European), Krongo (Kordofanian), Mwera (Benue-Congo), Tem (Gur), Mano (Mande), Tojolabal (Mayan), Cantonese (Sino-Tibetan), Cocamo (Tupi), Maung (Australian), Atchin (Oceanic), Abipon (Ge-Pano-Carib), Bari (Eastern Sudanic), Zuni (Native American isolate), and Nung (Tibeto-Burman).[1]

The next section contains a more detailed account of how diachrony in general and grammaticalization in particular provide a framework for tracking the similarities and differences among languages.

11.4 Grammaticalization paths as 'strange attractors'

In the literature on complex adaptive systems, 'attractor' is the name given to the path that a dynamic system takes. In a closed system, where there is no new input of energy, a fixed point can be the attractor, as in the case of a bob swinging on a string, which will come to rest at a fixed point or attractor. In a complex non-linear system, such as language, where there is new input of energy provided by language use, it can be observed that no cycle ever follows exactly the same path. Thus while cycles may seem very similar, yielding a global pattern, the details are always somewhat different (Larsen-Freeman 1997). Such a situation can be observed in language change, especially in grammaticalization, where paths of change can be identified that are cross-linguistically very similar, yielding a global pattern, even though the details show differences. In this section we will examine the paths of change for future markers across languages, demonstrating both similarities and differences. As mentioned earlier, viewing cross-linguistic patterns as patterns of change rather than as patterns of fixed states provides us with a more insightful basis of comparison.

In Chapter 10, the polysemy in the English future marker *will* was discussed and it was shown that *will* can indicate intention, prediction and willingness. That discussion made it clear that some of these meanings were expressed by *will* and not by the other English future markers. Also the occurrence of these meanings in specific constructions would suggest that they are particular to English. These facts would seem to make it very difficult to compare a grammatical category such as 'future' across languages. However, if we consider the way that grammatical morphemes develop diachronically, we find that the paths of development as well as the mechanisms behind them are very similar, providing us with a means for cross-linguistic comparison. The similarity among futures across languages can be summarized in the following set of grammaticalization paths proposed on the basis of documented changes as well as cross-linguistic synchronic patterns of polysemy (Bybee, Pagliuca and Perkins 1991, Bybee et al. 1994):

(1) Common paths of change resulting in future markers

'movement towards a goal'				
'desire, volition'	> intention	> prediction	>	epistemic or
'obligation'				subordinating
				modality

Future markers are defined as any grammatical marker that indicates a prediction made by the speaker (Bybee and Pagliuca 1987, Bybee, Pagliuca and Perkins 1991). In the cross-linguistic survey cited here, all other uses of markers used for prediction were taken into account.

The list of meanings on the left side are meanings of lexical sources that appear in constructions that come to express the future. Surveys of many languages reveal that it is common to find a stage at which intention is expressed. The intention use comes about through an implication from the original lexical meanings used in first person contexts where expressions such as *I want to, I have to* and *I'm going to* can reasonably lead to the inference that the speaker intends to do something. It sets the stage for the move to future (prediction).

Since new meanings arise in specific contexts, they do not immediately replace old meanings; rather there can be long periods of overlap or polysemy where old and new meanings coexist. This fact is a major factor in accounting for the oft-cited modal meaning of future markers. That is, meanings of futures that indicate volition or willingness (as found for instance in Danish, Nimboran, Bongu, Dakota and Tok Pisin [Bybee, et al. 1994: 254]) probably are retentions from the original meaning of the lexical items and constructions that grammaticalized to form the future. The same can be said for futures that also have uses indicating obligation (such as found in Inuit, Basque, Danish and Slave [Bybee, et al. 1994: 258]). Futures derived from constructions indicating movement towards a goal usually do not have modal uses except for expressing

intention. Thus a comparison based on grammaticalization paths affords a means to explain both similarities and differences among future markers.

Once the intention use is common in a language, the stage is set for the further inference of prediction, which can be taken to be the main diagnostic for a future marker. In the following examples from Chapter 10 *will* and *be going to* can be interpreted as expressing the intention of the subject of the sentence, a prediction by the speaker or just as likely, both.

(2) Madam, my lord *will* go away to-night; A very serious business calls on him. (*All's Well That Ends Well*, II. 4)
(3) She's *going to* take a poll in 1991 to find out what her chances are. (COCA 1990)

At each of the three stages mentioned so far – lexical source construction, intention and prediction – the new meaning is applicable in a wider range of contexts and thus the frequency of use increases at each stage. The generality of the meaning is also increasing and continues to increase. Later developments suggest that the meaning becomes applicable in a variety of contexts. In fact, the predictability of the path breaks down some in late stages and greater cross-linguistic diversity can be observed, as futures that have traversed this far on the path might be found to express imperative, the epistemic modalities of probability or possibility and to occur in some subordinate clauses (without their lexical meaning) such as protases of conditional sentences, temporal clauses and complements to certain kinds of verbs (Bybee et al. 1994).

In this way the grammaticalization paths in (1) act as 'strange attractors' in a complex adaptive system. We can observe this path being manifest in language after language and at different time periods, yet there are differences in detail from one manifestation of the path to another. One source of cross-linguistic differences are possibilities of differences in lexical source. Here it should also be mentioned that there are a few more attested lexical sources for futures that are less common, that is, temporal adverbs with meanings such as 'then', 'afterwards' and 'soon' and modals indicating ability and attempt. Interestingly, these very likely have gone through stages similar to the more common lines of development shown in (1), starting with 'intention'.

Another source of cross-linguistic differences in the manifestation of the future path is the existence of other constructions in the same functional domain in the same language. For instance, the existence of *shall* as a future marker earlier on limited the spread of *will*, which was not used with first person until recently (Coates 1983). In addition, as we saw Chapter 8, particular instances of constructions can be conventionalized, encouraging or inhibiting the spread of a construction and thus creating language-specific characteristics (Poplack, to appear).

The underlying mechanisms of change that occur in grammaticalization are the same in all languages, which leads to the similarity in the paths or the

attractors, especially insofar as these mechanisms are nested in cross-culturally common discourse needs and communicative contexts. These mechanisms discussed in Chapter 6 include generalization of meaning, habituation and pragmatic inferencing. Interestingly, even the inferences that push the changes forward seem to be very similar across languages and cultures (see section 11.6). However, the context for grammaticalization in each language may be slightly different; the cultural context may differ as well (see section 11.7). Thus no two paths of development will be exactly the same, though very similar paths will occur over and over again in the languages of the world.

The knowledge of diachrony, then, provides us with a way of comparing the uses of futures across languages and making predictions about what kind of development will come next. Casting these developments in a complex adaptive systems framework helps us evaluate both the similarities and the differences in paths of development.

Note that these facts about similarity among languages cannot be listed in an innate Universal Grammar. They are rather facts that show that the paths of development for grammar are similar in the same way that dunes of sand or waves on the ocean are similar: because the forces that create them are the same and these forces interact dynamically over time to produce emergent structures that are similar but never identical.

Within cognitive and functional linguistics, there is mounting agreement that when we search for the 'universals of language' we need to focus our search on the processes that create and maintain language structures, not on the structures themselves (see Givón 2002 and Verhagen 2002). For this reason, the current book has been focused on the processes that create linguistic structure. To review, we have been concerned with the effects of the chunking which occurs in sequential processing, as this provides for groupings of morphemes and words that underlie constructions and constituents. Categorization, the most basic of cognitive processes, establishes the units of language, their meaning and form. The Law of Contiguity or cross-modal association (James 1950, Ellis 1996) allows for symbolization or meaning-form associations. These processes in combination with the effects of repetition on memory and access provide us with an explanation for many of the properties of linguistic form. When we consider that the content with which languages deal – what people choose to talk about and how they choose to talk about it – and the social interactive situation are often similar, we have a concrete basis for understanding how and why all languages are alike. In addition, as we will see in the next section, this view of grammar provides a framework for working out a very plausible view of language origins.

11.5 Origins of language from domain-general abilities

After a long hiatus it has again become acceptable and even popular to speculate on the origins of language. Given what we know about language change,

particularly grammaticalization, which provides a well-documented account of how grammar emerges from repeated word sequences, there is every reason to suppose that the very first grammatical constructions emerged in the same way as those observed in more recent history (Bybee 1998a, Li 2002, Heine and Kuteva 2007). Moreover, the fact that we can relate the emergence of grammatical constructions to domain-general abilities that are present not only in humans, but also to varying degrees in other primates means that usage-based theory need not postulate an evolutionary event (whether of adaptation or mutation) by which the brain was rewired dramatically to contain the essence of Universal Grammar as necessitated in the theories of Pinker and Bloom 1990, Jackendoff 2002, Pinker 2003, and Chomsky 2006. Rather a theory based on domain-general cognitive abilities postulates the increasing capacity of such processes – increase in memory and access to memory, development of increasingly finer motor and perceptual skills, increased ability at imitation and at sequential processing, and greater abstraction in categorization, all of which could develop gradually while some form of language is being used (Bybee 1998a, Li 2002).

It is not my goal to pursue discussion of the biological foundations for language further in this section. Instead I will focus on theories of the evolution of grammar itself. The main thrust of this section will be to demonstrate that theories of the evolution of grammar must be based firmly on an understanding of language change and how it takes place.

In Chapter 6 I argued at length that the primary locus of language change is not in the first language acquisition process, but rather in the process of l anguage use. Unfortunately, many researchers have embarked on a study of language evolution subscribing to the erroneous assumption that language change occurs primarily as language is transmitted across generations (for some examples, see Briscoe 2003, Kirby and Christiansen 2003). Applying a Darwinian model to this view, replication would occur with each individual's acquisition of a grammar. Faulty replication would create a grammar that is not identical to the adult model and would therefore introduce change. In contrast, the usage-based view, as presented in Croft 2000, takes the replicator to be the linguistic element and replication to occur in each utterance produced by speakers in a community. As pointed out in previous chapters, innovations in utterances often involve small articulatory adjustments due to neuromotor accommodations or extensions of constructions to new but related contexts. Such changes, if they are repeated over multiple speech events, add up to recognizable changes in phonological and grammatical structure.

Thus the evolution of grammar from the usage-based perspective requires that cross-modal association is already possible; that is, language users have begun to associate sound with meaning. Then if two sound–meaning symbols (or words) are produced in sequence, the stage is set for the elaboration

of grammar, first through chunking, then through grammaticalization. The repetition of a two-word sequence can lead to the expansion of the lexicon through compounding, and compounds with repeated elements (*man-like, god-like, friend-like*) can lead to the development of derivational affixes (*manly, godly, friendly*). In addition, frequent word combinations can lead to the development of multi-word constructions and instances of such constructions can grammaticalize with repetition. It is important to note that innovations in the lexicon and the development of new grammatical elements and constructions through grammaticalization cannot occur in the first language acquisition process, but can only occur more gradually in language use.

Note that in this view, the first language or languages are thought not to be the same as present day languages. They would have had lexical items but not grammatical items or constructions. Grammar developed gradually as language was used and as the capacities of humans or our ancestors increased to accommodate a large vocabulary, more abstract categories and many automated sequences. Heine and Kuteva 2007 provide a set of explicit hypotheses, based on the now extensive literature on grammaticalization, which show how the modern categories of grammar could have been built up gradually in successive layers, as nouns and verbs followed their well-known grammaticalization paths. Thus rather than adopting a version of the uniformitarian hypothesis that says that the first languages had basically the same properties as documented languages, we should rather adopt the version that says that the processes of change were the same in the past as they are now (Heine and Kuteva 2007).

An alternate, but not mutually exclusive view, is taken by Wray and Kirby (Wray 2000, Kirby 2000, Wray and Grace 2007). These researchers do not take compositionality to be basic. Wray argues that complex structures could arise through the analysis of holistic structures rather than through the composition of simple structures. Wray 2000 notes the heavy use of holistic formulaic expressions in modern languages and Wray and Grace 2007 speculate further that unanalysed expressions may be even more common in social situations where one interacts primarily with people who share a similar background (see the next section for further discussion). Kirby 2000 demonstrates through a series of experiments that as the words of an artificial language are transmitted to new learners, the learners impose some order on the words, changing them to create recurrent parts that correspond to morphemes. Thus these researchers question the assumption made by many linguists, that compositionality – the regular and transparent combining of morphemes and words – is basic to grammar.

There is much to commend the view that holistic expressions are also natural. In Bybee 1985 I argued against the assumption that morphology is most natural when it is regular and compositional. By pointing to the fact that irregularity of form was most common in high-frequency items, I suggested that there is a natural place in the grammar for less analysable and more fused forms. The same

theme has appeared in the current work as well. The previous chapters have demonstrated that loss of analysability and compositionality and the increase in autonomy are the natural consequence of the way language is processed and indeed provide us with a source for grammar. However, those who propose a role for analysis of holistic expressions should not lose sight of the nature of linguistic change as documented over many languages and many centuries. While this record does show some cases of folk etymologies and backformations that indicate that holistic units have been analysed, the vast majority of grammar-shaping changes start with two or more elements and fuse them into one. Thus for every case of folk etymology, as when *hamburger* is analysed as consisting of the noun for a type of meat or other ingredient plus an element that must mean 'sandwich on a bun', giving us by analogy new words such *fishburger* and *veggie-burger*, there are hundreds, if not thousands, of documented cases of changes going in the other direction – from analysable, complex formations to unanalysable formations. Thus I would ascribe a minor role to the analysis of holistic expressions in the evolution of grammar and a major role to grammaticalization. Note also that giving the analysis of holistic expressions a major role in language evolution is based on the assumption not accepted here that linguistic change occurs principally in language transmission.

11.6 Social factors that shape grammar

It follows from the premise of this book – that linguistic structure emerges through language use – that the social and cultural context in which language is used would have an impact on the structures that are created. We have already seen that frequency or repetition leads to loss of analysability and compositionality, reduction of form, generalization of meaning and conventionalization of inferences. To the extent that the conditions under which language is used are similar across cultures, the substance and form of grammar will also be similar; to the extent that these conditions differ, languages may has grammars of different types. Thus we might expect to find differences in typology relating to some extent to differences in cultural context. In contrast, a theory that relies on a set of innate givens, such as Universal Grammar, has very restricted means by which to account for typological differences among languages. In this section a few factors concerning the social interactional contexts in which language is used are discussed to show how these may impact grammar, providing in some cases for similarities across languages and in other cases for differences.

11.6.1 Similarities in pragmatic inferencing

Throughout the discussion in previous chapters we have seen the important role that pragmatic inference plays in semantic change, particularly in

grammaticalization and the creation of new constructions. Clearly pragmatic inferencing is a universal mechanism that contributes to the creation of grammar. In section 11.4 we also noted that the similarity among semantic paths of grammaticalization across languages points to the fact that even across cultures that may be rather different, very similar inferences are made in similar situations. In the discussion of futures we saw that the inference of speaker intention is important in starting expressions of movement towards a goal, volition and obligation towards grammaticalization. A second inference of prediction also occurs to yield the future meaning. Since the same semantic sources and paths of change are documented across languages, it appears very likely that the same inferences are made in distinct cultures.

Another set of developments that point to inferences that are cross-linguistically similar are perfects and perfectives that can also be used to indicate present state. This occurs in Sango, Palaung, Tok Pisin, Engenni, Trukese, Island Carib, Kanuri, Mwera and in the Preterit-Present verbs of English (Bybee et al. 1994: 70–8). In these cases, when one expresses the concept of having entered a state, as in 'I learned' or 'it got dark' the inference is that the resulting state still holds: 'I learned' therefore 'I know'; 'it got dark' therefore 'it is dark'. Thus by the conventionalization of this implicature, polysemy results so that the perfect or perfective marker signals present with stative predicates. The impressive cross-linguistic similarity is a strong indication that people in different cultures can make very similar inferences.

A third example concerns the inference of causation from the expression of the temporal relation 'after'. Just as *since* in English has changed from having only a temporal meaning to expressing cause as well, this development has occurred in other languages. The English case, as discussed in Traugott and König 1991, shows a diachronic development from the temporal meaning in (4) to the cause meaning in (6) via examples such as (5) in which the context leads one to make a causal inference from the temporal meaning.

(4) I think you'll all be surprised to know that *since* we saw Barbara last, she made an amazing trip to China. (COCA 1990)
(5) After 50 years of sleepwalking, he hasn't walked once *since* he started taking the drug. (COCA 1990)
(6) *Since* the hunters all have CB radios, they can warn each other before he even gets close. (COCA 1990)

Heine and Kuteva 2002 cite such polysemy in English, French, Basque, and Aranda. The cross-linguistic pattern suggests that language users are particularly interested in finding causal relations even where they are not explicitly expressed. Thus we find cross-linguistic similarity in the actual inferences made: intention, prediction, resulting state and cause, and perhaps many others.

Of course, given different social and physical conditions, some inferences will undoubtedly vary across cultures.

11.6.2 Inferencing and morphological typology

Despite these similarities in the actual content of inferences, differences among languages also occur because of differences in the nature and extent of inferencing in discourse. In Bybee 1997 I examined the role of discourse inferencing in determining how far a language would carry the process of grammaticalization. The results of extensive cross-linguistic comparison of grammaticalization in Bybee et al. 1994 show that there are differences in the extent to which grammaticalization is carried out. We found in languages of the analytic or isolating type that not only were the grammaticalized forms longer and less fused with the verb (being less phonologically reduced in general), but also the meanings of grammatical categories were more specific and represented earlier stages of grammaticalization paths. For instance, a robust finding of Dahl 1985, Bybee and Dahl 1989 and Bybee et al. 1994 is that languages that lack inflections – that is, categories that are affixed and obligatory – also lacked perfective/imperfective and present/past distinctions. Such languages – the analytic types – tend rather to have perfects (or anteriors), which represent the earlier stages on the past and perfective path, or progressives, which represent the earlier stage of the present or imperfective paths.

This finding echoes the classification of morphological types proposed by Sapir 1921. While subsequent researchers tend to think of Sapir's proposed types as purely a matter of form, his actual discussion relates form to meaning and proposes that languages of different morphological types express different types of meaning. Thus Sapir distinguishes between Concrete Relational Concepts and Pure Relational Concepts, his names for types of grammatical meaning. He distinguishes these two types in terms of the degree of abstractness of their meanings. He does not place any grammatical categories permanently in one or the other category, but rather argues that a category such as number, gender or aspect may be more concrete in one language but more relational in another. For instance, where number is marked only on nouns it is more concrete, but where it also marks agreement on demonstratives, adjectives or verbs, it is more relational. In Bybee et al. 1994 and Bybee 1997 we proposed a rough equivalence of the more concrete relational concepts to meanings that occur earlier on grammaticalization paths and those that are more purely relational to more grammaticalized meanings. Given the parallelism of the development of form and meaning, then, languages which do not carry grammaticalization through to affixation would also not carry the semantic grammaticalization as far as inflectional languages do.

Thus the traditional morphological typology is underlyingly a typology of how far grammaticalization is carried in a language. The quantitative test

of this hypothesis in Bybee et al. 1994 was based on the formal properties of morphemes associated with verbs in a seventy-six language sample on the basis of which each language could be classified by the phonological length of its grammatical morphemes, their dependence upon surrounding material and their fusion with the verb. Using these formal measures, we tested the correspondence of the overall morphological type of a language with the degree of semantic grammaticalization of its markers of completive, resultative, anterior (perfect) perfective and past, listed here in order from least grammaticalized to most grammaticalized. The correspondence with dependence on surrounding material and fusion with the verb were highly significant – the more fusion and dependence in the language in general, the more likely that the language had a highly grammaticalized morpheme for perfective and past. No correspondence was found for the length of the morphemes in the language. Of course, this is to be expected as no one has ever proposed a morphological typology of language based solely on the length of grammatical morphemes.

Thus the hypothesis that morphological typology depends upon how far a language carries the grammaticalization process is supported. On the inflectional end of the scale, grammaticalization proceeds to the development of the most abstract and general meanings. These are expressed by affixation and in some cases, because of further phonological changes, by stem changes. On the analytic end of the scale, grammaticalization proceeds less far – grammatical morphemes do not become affixes, nor do their meaning changes proceed so far as to establish the most abstract and obligatory categories of meaning. Rather, it appears that grammatical morphemes are replaced by other newly grammaticalizing morphemes before the older ones have had a chance to reach the end of a path of development (Lin 1991).

What prevents grammaticalization from proceeding as far in some languages as it does in others? In Bybee 1997 I suggested that an essential process in the later stages of grammaticalization is not available in languages of the analytic type. This process involves a particular type of inferencing which makes a category become obligatory, as obligatoriness of categories is the defining feature of inflection.

A characteristic of analytic or isolating languages is the lack of obligatory expression of items such as pronouns and the lack of obligatory categories, defined as categories for which some exponent must appear in the phrase or clause. Consider the Chinese sentence used by Bisang 2004 to illustrate this property of analytic languages:

(7) wǒ bú jiàn tā, yǐ shǐ sān shǐ duō nián; jīntiān ø jiàn ø le.
I NEG see he already be 30 more year; today see PF
'I haven't seen him for more than 30 years. Today [I] saw [him].'

Note that in the second clause there is no need for the expression of the pronominal forms as they can be inferred from context. Note also that there is no expression of tense in the first clause, but the lexical expression '30 years' makes the temporal reference clear. Typically, the form of expression in analytic languages contains very little in the way of redundancy or repetition. For instance, there are no explicit grammatical markers of the role of the arguments in the clause. Given a certain flexibility in word order, the listener is left to infer the relations among the NPs in a clause. In such cases, the semantics of the NPs, along with real world knowledge of agentivity, is the most important guide to semantic roles (Li, Bates and MacWhinney 1993). Rather than relying on explicit grammatical markers or word order, the listener must work actively to infer the relations among the NPs that the speaker intends. Similarly, in the domain of tense and aspect a lot can be left unexpressed; the listener again must apply the most reasonable inferences.

Bisang 2004 points out that there is a high degree of indeterminateness in both the morphosyntax and the lexicon of the analytic languages of East and Southeast Asia. Grammatical markers in these languages are polysemous and can express meanings from various functional domains, depending upon the context. Lexical items may be interpreted as either nouns or verbs, again, depending upon the context. As a result, grammatical markers lack two properties that Bisang considers would otherwise lead to the development of obligatory categories: frequency of use and a clear-cut semantic domain. Because markers are not used redundantly, they do not undergo the kind of frequency increase that usually characterizes grammaticalization. Because they operate in various semantic domains, no paradigms emerge. These two properties, which both reference the use of markers in discourse context, are features of these languages which inhibit grammaticalization.

Let us now consider the role of redundancy in promoting grammaticalization. Redundant expression can be of at least two types. One sort of redundancy comes about when a speaker expresses an idea as part of an utterance, where that particular idea would be assumed even without expression. For instance, English uses modal elements expressing obligation much more often than the cognate or similar items would be used in other European languages. For instance, an American English speaker would say *I have to go now* in the same context in which a Dutch speaker would simply say *Ik ga nu* or a Spanish speaker would say *me voy ahora.* If the context is one of, say, going to a doctor's appointment, the notion of obligation is implicit; however, English expresses it and the other languages do not. The increase in frequency in early stages of grammaticalization are probably due to this sort of redundancy – where a notion might have gone unexpressed in the past (because it was easily inferable), it is now expressed wherever it is intended.

A second type of redundancy is supplied by the actual linguistic elements: within a discourse in most cases one expression of tense may be enough if

several clauses have the same temporal reference. However, in languages in which tense is obligatory, it is expressed in every clause. Similarly, in languages with subject-verb agreement it appears whether it is necessary for comprehension or not. The same type of redundancy occurs where determiners and adjectives agree with the noun in number and gender. This second type of redundancy indicates an even more advanced stage of grammaticalization, the stage at which categories have become obligatory. Both types of redundancy are characteristic of synthetic languages but not analytic languages.

What leads to the development of redundancy and obligatoriness? These are of course very difficult questions. In the extreme frequency increases during grammaticalization one senses an inexorable movement by which each frequency increase leads to another. One possible factor is that the constructions that are grammaticalizing become all the more accessible because of their frequency; not only is their articulation automated, but their cognitive access is as well. One might say they reach a high level of resting activation and as a result are more likely to be selected for production.

Redundant activation might be inhibited, however, by discourse conventions that favour non-redundant utterances. In such cases, repeated constructions must be interpreted as new contributions to information. Given such interpretational conventions, speakers are not likely to use constructions redundantly. One source of increased frequency is thus constrained.

As mentioned in Chapter 10, high frequency of use, including redundant use, is a prerequisite to obligatoriness. As suggested by García and van Putte 1989, obligatoriness arises by pragmatic inference. If the expression of a category becomes common enough, the listener is entitled to infer that if the category is NOT expressed, then its opposite was intended. Thus the absence of expression comes to be considered the zero expression of the complementary category. In section 10.7 this development was illustrated with the English Simple Present, which developed an habitual interpretation when the Progressive grammaticalized.

Consider now conventions for making inferences. In a culture in which utterances contain fewer redundancies, each element is taken as meaningful, so the absence of elements can mean either that the absent meanings are intended to be inferred, or that they are not intended. Thus the listener is required to fill in information that goes unexpressed. The listener is not accustomed to make the kind of inference that assigns meaning to the absence of mention. Compared to synthetic languages with many obligatory categories, the listener in an analytic language does not have linguistic cues that eliminate certain possible meanings, as when a case marker or verb agreement confirms which NP is the subject. Rather the listener is making probabilistic judgements based on semantics and prior context to determine the role of NPs, as well as other factors, such as temporal reference. With this type of inference,

obligatory categories will not become established; thus an analytic language will remain analytic as long as the inferencing strategy remains the same.

It is important to note that inferencing strategies are conventional and must be learned. Children learn through experience with utterances in context what can be inferred and what must be explicitly expressed, and as we have seen this differs from language to language. Once such conventions are established, I would argue, they have an effect on how far grammaticalization can proceed in a language. Are such cultural conventions related to other properties of the culture? I see no reason to assume such relations as far as inferencing strategies are concerned. However, as we see in the next section, certain types of morphological categories may be highly related to the nature of the culture within which the language is spoken.

11.7 Deictic morphology and cultural type

It has been noted that speech used in an intimate environment – where participants know each other well and share many experiences – is different from that spoken in more public contexts, among participants who are not intimate and who cannot be presumed to share many past experiences or current conditions (Bernstein 1972, Kay 1977, Givón 1979, Perkins 1992). In situations where speakers share backgrounds, utterances can, for instance, have more pronominal use or omission of NPs, fewer subordinate clauses and the markers of those clauses. A number of researchers have noticed similar differences between spoken and written language (Chafe 1982, Biber 1986). Givón 1979: 207–33 writes of a 'pragmatic' mode characteristic of unplanned and informal discourse compared to a 'syntactic' mode used in more planned and formal discourse. He argues that evolutionarily the pragmatic mode precedes the syntactic one, which develops as a response to the speech situation in a more complex culture where we often talk to strangers.

Perkins 1992 has devised a rigorous means of testing the hypothesis that the social and cultural context in which language is used affects grammatical structure. In particular, working from the observations referred to above from Bernstein, Givón and others, Perkins hypothesizes that languages spoken in cultures where small groups share a limited physical and social background will have more inflectional or affixal markers of deixis than languages spoken in cultures where large numbers of people of diverse background communicate with one another. The hypothesis relies on the fact that inflectional affixes arise by grammaticalization and that in order for forms to grammaticalize they must be used with high frequency. Thus Perkins proposes that in cultures where communication commonly occurs among familiars, deictic expressions, such as *here, there, now, then, she, he* will occur often enough to become grammaticalized. In contrast, in cultures in which communication has to be more

explicit, such markers will not grammaticalize as readily. Because there is always the cyclic loss and replacement of grammaticalized forms, as cultures grow more complex and the speech situation changes, deictic inflections will be lost and not replaced.

The test of this hypothesis relies on resolving three important issues.

First and foremost, the hypothesis must be tested on a large sample of cultures and languages, but it is very important that genetic and areal bias in such a sample be controlled. Perkins solved this problem by using a sampling technique by which he chose languages randomly from a matrix that separated languages by genetic affiliations and the potential for areal contact. It was important that the languages be chosen randomly rather than 'by convenience' so that additional bias did not creep into the sample, as only well-studied languages came to the fore. Perkins' selection by this method yielded a sample of forty-nine languages on which to base his study.

Second, a method of measuring cultural complexity must be selected. Perkins used a scale derived from the report in Murdock's *Ethnographic Atlas* (1967 and 1967–71) based on nine cultural features which referenced the type and intensity of agriculture in a culture, rules for inheritance, regional organization, craft specialization, class stratification and size of settlements. Such measures are appropriate for the hypothesis as they indicate the extent to which members of the society share background assumptions and current presuppositions.

Third, for the linguistic test of the hypothesis, Perkins selected inflectional markers of deixis, which include person markers bound to nouns or verbs, dual marking (usually in second person), the inclusive/exclusive distinction in first person, bound demonstratives, and inflectional tense. In addition, Perkins coded gender distinctions in person markers as a frequently occurring, but non-deictic category in order to test whether the absence of deixis is merely due to an absence of inflection.

The results of the survey of the forty-nine languages/cultures supported the hypothesis. A significant correspondence was found between person affixes on nouns and verbs and cultural complexity such that languages spoken in the less complex cultures had more person affixes. A significant correspondence in the same direction was found for the dual distinction and the inclusive/exclusive distinction. The presence of tense affixes on verbs showed a trend in the predicted direction and came close but did not reach significance. The few cases of demonstrative elements in nouns and verbs also showed a non-significant trend in the predicted direction. Thus most of the categories tested aligned with cultural complexity in the predicted way. In contrast, the non-deictic category tested, gender agreement on verbs, showed a non-significant association with cultural complexity (in the same direction as the deictic affixes), indicating that it is not just the presence of inflection that is predicted by cultural measures, but deictic inflection in particular, as predicted by the hypothesis.

Other attempts at similar hypotheses have been less successful. It is not known how they would fare if tested empirically because the proponents of these theories have not submitted them to empirical testing. The prospect for empirical success, however, is diminished by certain flaws in the reasoning behind them. These hypotheses are based on the observation that second-language learners simplify aspects of grammar, particularly the inflection (Trudgill 2001, Wray and Grace 2007). Extreme examples of simplification occur in pidgin and creole languages. Such languages have fewer inflectional categories than languages with more normal development and they have fewer inflectional categories than their lexifier languages (the languages from which the majority of the vocabulary is derived) (as noted by many researchers, e.g. Bickerton 1981, McWhorter 2001). It is also known that adult second-language learners often do not fully master the inflectional system of the target language and on this basis, Wray and Grace 2007 advance the opinion that languages used 'exoterically' – that is, when talking to strangers – would tend to lose morphological distinctions. Wray and Grace 2007: 551 state the hypothesis as follows:

> Thus, languages that are customarily used exoterically will tend to develop and maintain features that are logical, transparent, phonologically simple and, significantly, learnable by adults. (Thurston 1989; Trudgill, 1989, 2002)

These researchers assume that the mechanism for the loss of inflection is the process of second language learning. Thus they claim that languages which are often the target of second language learning will develop properties that simplify them and make them more learnable by adults.

Several problems come to light when this hypothesis and its associated mechanisms are examined carefully. We will examine some of these problems here.

First, it is not necessarily valid to assume a continuum from what happens in the pidginization process to what happens in language contact situations in languages undergoing normal development (McWhorter 2001, Dahl 2001 contra Trudgill 2001). As is well known, pidgin languages arise in restricted social situations (plantations, trade situations) where multiple languages are spoken natively. A particular language is chosen for communication in this setting, but access to native speakers of the language, and thus the language itself is very limited. Failure of adult learners to master the language is at least in part due to this limited access. In contrast, in more ordinary cases of language contact or bilingualism (as in the case of immigrant populations, such as guest workers in Europe), second-language learners are embedded in the target language and culture. While in such cases adults still show a lesser ability to acquire a language than children, their failures do not affect the language as a whole. Rather, the effect of second-language learners is fleeting in the sense that their children have full access to the language and acquire it, becoming native

speakers. The presence of adult second language learners does not change the language; rather the immigrant population gradually shifts to the language of the majority.

Second, it is important to note that not all inflection is lost in pidgin and creole languages. Roberts and Bresnan 2008 survey the categories lost and retained in twenty-seven pidgin languages from all around the world and report that fifteen of these languages retain some inflection. Certain tendencies for which type of inflection is retained also emerged:

> We have encountered evidence that the reduction of inflection is asymmetric and not always total. Inflections that contribute semantic and grammatical information pertaining to the stem are retained slightly but significantly more often than inflections that pertain more to building the syntax of the sentence outside the word. (Roberts and Bresnan 2008: 293)

Nor do second-language learners eliminate all morphology from their version of the target language (Dietrich, Klein and Noyau 1995, Prévost and White 2000). Studies of adult second language learners in natural settings (outside the classroom) are inconclusive on the issue of the use of inflections. Dietrich, Klein and Noyau 1995 and Klein and Perdue 1997 argue that the Basic Variety of the first 30 months shows very little morphological marking of tense or aspect; however, after this period some learners go on to use some inflections for tense (especially if the target language is French or English). As for agreement, Prévost and White note many correct agreement uses in the first three years for learners of French and German. Thus it is certainly incorrect to conclude that adult language learners eliminate all inflection. It is also incorrect to assert, as Wray and Grace do, that adults learn rules, while children learn the specifics and generalize less. Dietrich, Klein and Noyau observe that for all the target languages represented in their study (English, German, Dutch, Swedish and French) adult learners started with the irregular past tense formations, apparently overlooking the simpler rules of the regulars.

Finally, theories based on the notion that language change occurs in transmission to new speakers and in particular that second language adult learners simplify inflection provide no means for explaining why languages have inflection in the first place. In contrast, the theory of Perkins, which is based on the well-supported premise that language change occurs in language use, explains through grammaticalization why languages have inflections in the first place, as well as why deictic categories are not replaced in certain cultural contexts. With respect to grammaticalization, it is important to note that when inflectional categories are created anew in pidgin and creole languages, the process by which this happens is the same as in languages with normal development – that is, new categories are created by grammaticalization. This fact provides further evidence that the usage-based process of grammaticalization is responsible for

the origins of grammar wherever grammar is created – in language origins, in pidgin and creole languages as well as in mature languages.

11.8 Grammaticality, frequency and universals

In this final section we treat another way in which usage patterns determine cross-linguistic patterns by considering the factors that make particular constructions frequent or infrequent within a culture. An important component of this discussion rests on the usage-based notion that high frequency of usage leads to conventionalization and further elaboration, while very low frequency of use leads to unacceptability and eventual loss. Thus we find that some construction types are robustly represented across and within languages (transitive clauses, possessive constructions) while others vary considerably across languages in their frequency of use (serial verbs [Hopper 2008]), and some that are rare in the languages in which they occur are ungrammatical in others (oblique relative clauses [Keenan 1975]; see discussion below).

In usage-based theory, grammaticality or acceptability judgements are considered to be gradient; both grammatical and ungrammatical combinations of words, morphemes or sounds can be rated for degrees of acceptability. As mentioned in Chapter 5, acceptability judgements within a language are postulated to be based on familiarity, where familiarity rests on two factors: the frequency of a word, construction or specific phrase, and similarity to existing words, constructions or phrases. Items will be judged as acceptable to the extent that they are frequent in the subject's experience or similar to frequent items. In the experiment reported in Bybee and Eddington 2006, the stimuli were taken from corpora and thus were presumably grammatical, but subjects were still able to rate them for degrees of acceptability. The highly significant results showed that high-frequency verb + adjective combinations were judged most acceptable followed closely by lower frequency combinations which were semantically similar to high-frequency combinations. Low-frequency combinations lacking semantic similarity to high-frequency ones were rated the least acceptable.[2] Thus we view the line between extremely low frequency and ungrammaticality as a gradient one.

The same factors that make a construction frequent or infrequent in one language can make it completely acceptable or unacceptable in another. High-frequency patterns are highly conventionalized and may be highly productive, while rare patterns may be retained only in fixed phrases or fall into unacceptability. Hawkins (1994, 2004, 2009) proposes the Performance-Grammar Correspondence Hypothesis, which he states as follows:

> Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments. (Hawkins 2004: 3)

Hawkins envisions the primary factor determining frequency in corpora to be ease of processing, but it should be noted that a wide variety of factors influences frequency in corpora. In the following I would like to mention some of these and demonstrate how they influence both frequency or infrequency in language use and the patterns of occurrence in the languages of the world.

11.8.1 What people want to talk about

We have already seen that the most usual inferences or assumptions can determine factors such as the distribution of zero expression, as zeroes take on the meaning that is the most usual in the context. Highly generalized grammatical categories also gravitate towards what people talk about most – perfectives for narration, present for habitual states and situations. In addition, one might mention the high frequency of first-person-singular pronouns and verb forms, as conversation is highly subjective. First-person-singular verb forms are often highly autonomous, resisting change. Thus many cross-linguistic generalizations about both form and meaning are partially determined by what people tend to talk about.

11.8.2 What constructions speakers choose to use

There are also strong tendencies in the particular ways that information is presented and interaction is managed that determine some cross-linguistic properties of grammar. The tendency to put topics first and to choose (human) agents as topics leads to the development of the category of 'subject' and its tendency to occur before the object (Tomlin 1986, Siewierska 2002). Hawkins 1994, 2009 reports that in languages which allow both the order subject-object and object-subject, the former is much more frequent in discourse; this is paralleled, of course, by the cross-linguistic finding that subject-object order is much more common than the opposite.

Strategies for organizing discourse can also lead to the establishment of grammatical properties such as the inflectional marking of argument roles on verbs. Du Bois 1985, 1987 shows that a persistent discourse strategy in Sacapultec Maya introduces full NPs into a narrative in the absolutive role, usually as the subject of an intransitive. Further narrative mention of the referent then occurs in the ergative, but this reference is signalled only by agreement marking on the verb. Du Bois takes this pattern to be the source of the zero expression of the third singular of the absolutive – it co-occurs more often with a lexical NP, while the ergative inflection derived presumably from a pronoun and thus has overt marking.

The NP Accessibility Hierarchy was one of the first widely discussed cases in which it could be shown that what was rare in one language was non-occurring

(or unacceptable) in another (Keenan 1975). This hierarchy is based on the role of the NP within the relative clause that is co-referential with the head noun. Keenan and Comrie 1977 demonstrated on a large cross-linguistic sample that if a language could form a relative clause on a case role in the following list, it could also form a relative clause using all the case roles to the left of that one.

Accessibility Hierarchy:
Subject > Direct Obj. > Indirect Obj. > Oblique > Genitive > Obj. of Comp.

That is, some languages allow relative clauses to be formed only with Subjects (Malagasy), others only on Subjects and Direct Objects (Welsh), and so on.[3] Keenan 1975 also demonstrates that in English written prose, the frequency of occurrence of each type of relative also follows this scale: subject relatives are the most common (constituting 46 per cent of the set of 2,200 relative clauses examined), direct object relatives the next most common (24 per cent), oblique and indirect object relatives next (15 per cent) and genitive relatives last (5 per cent). This correlation between acceptability in the languages of the world and frequency within one language is open to various interpretations. Keenan 1975 suggests that "there may be some sense in which it is 'easier' or more 'natural' to form RCs on the Subjects (or higher) end of the [hierarchy] than on the lower end" (1975: 138). Keenan and Comrie 1977 argue that relative clause formation with NPs in certain grammatical roles is psychologically easier because the meanings are easier to encode. They also cite studies in which it is shown that children comprehend relative clauses formed on the left end of the scale more easily than those on the right. Diessel and Tomasello 2005 offer a competing explanation of the ease with which English-speaking children use subject relatives: a subject relative retains the same word order structure as in main clauses.

Hawkins 1994 refers to 'complexity' in his suggested explanation. He proposes a formal grammatical account in which the structural description of each grammatical role is compared, and it is found that one cannot supply a structural description for a Direct Object without making reference to the Subject; further Indirect Objects require reference to Subject and Direct Object and so on. From these characterizations, one would have to further say that the 'easier', 'more natural' or less complex structures occur more frequently in the discourse of a single language and are more likely to be acceptable across languages; see Hawkins' Performance-Grammar Correspondence Hypothesis, given in the previous section. The causal link between the intra-language and the cross-language hierarchies must be that what is rarely used may come to be considered unacceptable.

The explanations provided by Keenan and Comrie and Hawkins leave much to be desired. Keenan does not specify what he means by 'easier' or more 'natural' nor does he say why certain meanings are 'easier' than others. Hawkins' formal proposal is strictly internal to the grammar in that it takes as givens

the notions of Subject, Direct Object, and so on, notions that themselves need explanations.

Another possible explanation comes from what is known about how speakers tend to organize their discourse. Fox 1987 and Thompson and Fox 1990 take this approach to the examination of relative clauses in English conversation. Of course, relative clauses may not be used in exactly the same way across languages, but their studies of the use of relative clauses in English conversation is strongly suggestive of a discourse-based explanation for the Accessibility Hierarchy. Fox 1987 notes in the data she examined that subject and object relatives were equally frequent, but that most of the subject relatives were actually not agents but rather the subject of an intransitive verb. Thus by far the preponderance of examples shows the noun phrase playing the role of absolutive – subject of the intransitive or object of the transitive in the relative clause. In Fox's study and a more detailed one by Thompson and Fox 1990, it is determined that the discourse or conversational role of relative clauses is to establish the referent as relevant or related to referents that are already given in the discourse. This is done by presenting relevant characteristics of the referent (the subject function, as in 10) or by presenting the referent as the object of a transitive predication in which the agent is a pronoun (one of the discourse participants or referents already introduced, as in 11). Not that other functions are not possible; indeed they are, but these functions performed by relative clauses are the most common, making the grammar of subject and object (or absolutive) relatives the most accessible and conventionalized.

(10) She's married to this guy who's really very quiet.
(11) This man who I have for linguistics is really too much.

If other languages also use absolutive relatives for these functions, they will be frequent in discourse, and other types will be infrequent and may even be ungrammatical. Thus there are languages such as Dyirbal (Dixon 1972) in which only absolutives can be relativized and agents (ergatives) cannot.

For the Accessibility Hierarchy, then, we see various explanations being offered: processing preferences, semantic ease/difficulty, grammatical complexity, and discourse functions. In each case, the link between these factors and the relative grammaticality or acceptability of the structure is frequency in usage. One could even argue that none of these proposed explanations is necessary since the higher frequency of absolutives (subjects and objects) over other argument types would in itself render them more likely to be relativized and thus more acceptable.

The question of why more relativization positions are acceptable in some languages than in others can be addressed in the context of Perkins' theory of communication in more or less intimate communities. As outlined above,

reference can be established through the use of deictic markers in smaller, more intimate cultures because of shared knowledge. In larger, more complex cultures, however, more explicit means for establishing reference, such as relative clauses, provide a necessary strategy for making reference explicit. Thus Perkins hypothesized a relation between a scale of cultural complexity and the extent to which a language allows relative clauses on the right end of the Accessibility Hierarchy. Using the languages whose relativization possibilities were discussed in Keenan and Comrie 1977, Perkins established that this cultural to grammatical association was significant, further supporting the hypothesis that contexts of language use determine what structures are grammatical.

11.8.3 Specificity or generality of meaning

Both the specificity and the generality of meaning have an effect on what is frequent in a language. Specific members of grammatical categories, such a dual number, occur less frequently within a language (Greenberg 1966) and are also more prone to loss than singular and plural. In contrast, as grammaticalizing elements, such as verbs, become more frequent and more generalized in meaning in the grammaticalizing construction, they may lose the ability to occur as main verbs. Thus the English modal verbs, such as the ancestors of *can, may, shall, will* and *must*, occurred in Old English primarily in their finite forms, with infinitive and gerund forms being very rare or non-existent. Many dialects of English now find the use of these auxiliaries as main verbs to be unacceptable, as seen in the double modal examples such as *shall can*, which formerly occurred but which are no longer acceptable. Thus the fact that many auxiliaries in the languages of the world lack nonfinite forms is due to the extreme bleaching of their meaning.

11.8.4 Processing ease/difficulty

As mentioned above, Hawkins 2004 attributes high or low frequency in discourse to processing ease or difficulty. One example that illustrates this point concerns the tendency for members of the same constituent to be adjacent syntactically. Example (12) shows the complement of *waited* to be adjacent to that verb, while in example (13) it is not.

(12) The man waited for his son in the cold but not unpleasant wind.
(13) The main waited in the cold but not unpleasant wind for his son.

Hawkins shows that tokens such as (12), which are much easier to process, are also more common in corpora of English, and that examples such as (13) would be unacceptable in some languages (Tomlin 1986).

11.8.5 Conventionalization of thinking for speaking

Another determinant of what is frequent in languages is conventionalized ways of packaging concepts for speaking, or 'thinking for speaking', as Slobin (1996) puts it. Slobin (1997a, 2003) deals primarily with the lexical features of verbs and what features they incorporate. He finds that languages fall into different types, depending upon whether they tend to incorporate directional information into motion verbs, as in Spanish, where verbs such as *entrar* 'go in', *salir* 'go out', *bajar* 'go down' and *subir* 'go up' are good examples, or whether they tend to include manner of motion information as in English *amble, saunter, run* or *swim.*

Another more grammatical example of thinking for speaking might be the extent to which a language uses serial verb constructions, that is, constructions in which two or more finite verbs are used inside the same clause to form part of the same predication (Hopper 2008: 254). Languages in which such constructions are quite commonly used are located in West Africa, Papua New Guinea and in other places throughout Africa and Asia. Serial verb constructions are not common in European languages, but they are not unknown. For instance, Hopper 2008 presents an in-depth analysis of the *take + NP and* construction of English, as in the following sentence:

(14) And unfortunately we are going to have to *take all these people and* squish them into a church that seats four hundred…

This example represents a serial verb usage in the sense that *take* and *squish* do not represent distinct events, but rather together provide the predication. Other examples in English are *go get,* as in *Let's go get some coffee,* and the *try and* construction.

Hopper's (2008) point is that what may be a minor construction-type in one language may be a dominant one in another language. That is, it seems all languages are capable of arriving at serial verb constructions through grammaticalization, but only some languages carry the tendency to an extreme. Surely this is not because one set of speakers needs serial verbs more than others. Rather, it seems that a convention of thinking for speaking – packaging information in a certain way – may become established, and then extend to more and more verbal sequences in a language.

11.9 Conclusion: explanation for linguistic structures

Following the Greenbergian approach, which we have seen presages the complex adaptive systems approach, we can consider similarities and differences among languages at various levels. At the level of specific constructions, inventories, or lexical items we tend to find some core similarities of both form and

function, but with many differences. Croft 2001 argues that constructions are necessarily language-specific, yet in particular domains such as voice one can categorize constructions and find similarities among them on several dimensions, including both grammatical and distributional properties. These similarities are related to the diachronic sources from which the constructions arise and how advanced they are on their particular grammaticalization paths.

To take an example we have discussed previously, one can give a semantic definition to a construction that expresses future tense by saying that one of its uses should be for a prediction by the speaker. This definition would specify a core set of constructions cross-linguistically but they would differ in many ways: some might also express other meanings, such as intention, obligation, willingness or probability. Some might be more frequent in discourse than others; some might be inflectional and others periphrastic. Some might be prohibited from *if* clauses while others are allowed there. As we have said before, to understand the differences we can trace their diachronic development: the particular lexical source for the construction will determine which modality meanings – obligation, volition or willingness – occur; the extent of the development along the path will determine the relative frequency of modality, intention and prediction readings as well as the formal properties of the marker.

Thus the paths of change for constructions – such as voice, tense and aspect constructions – project stronger universals than the simple cross-linguistic comparisons of synchronic states. However, these cross-linguistic paths can also be further decomposed into the mechanisms and factors that create them as language is used. As mentioned in Chapters 6 and 10, chunking and phonological reduction, along with meaning changes traceable to habituation, generalization and inferencing, give rise to these changes. Thus an even stronger level for the statement of universals resides in the mechanisms of change that produce the paths as these do not vary across languages or across time (Bybee et al. 1994, Bybee 2001a, 2006b).

Yet another dimension in which language universals can be identified is constituted by the continua we have identified in the properties of constructions: analysability, compositionality, autonomy, schematicity, productivity and prototype effects in categories. All constructions in all languages have these properties to some degree or other. Thus while grammar itself is emergent and language-specific, the properties of the units of grammar on these dimensions are quite comparable across languages.

As we have seen, however, even these properties derive from the more basic cognitive processes of categorization by similarity, chunking of repeated sequences and association by contiguity. Categorization by similarity produces the categories of meaning of words and constructions, the grouping of bits of experience into the formal units of language, the categories for slots in

constructions and degrees of analysability. Chunking of repeated sequences of units cements the parts of constructions together, and gives us degrees of constituency or coherence among morphemes and words. Association by contiguity allows forms to take on meaning and allows meaning to change from association with context and with frequently made inferences.

These domain-general processes operate through repetition on a massive scale, within individuals and certainly within communities; this repetition within the context of what humans like to talk about and how they structure their discourse gives shape to the grammar and lexicon of particular languages. To the extent that context, meaning, and discourse patterns are shared across languages, similarities in structure arise. Thus taking language to be an embodied activity that occurs in real time, in real situations and passes through real cognitive systems has great potential for leading to the explanation of what we perceive as linguistic structure.