JOAN BYBEE

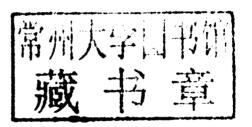# Language, Usage and Cognition

Language, Usage and Cognition

Joan Bybee
University of New Mexico

CAMBRIDGE
UNIVERSITY PRESS

Language, Usage and Cognition

Language demonstrates structure while at the same time showing considerable variation at all levels: languages differ from one another while still being shaped by the same principles; utterances within a language differ from one another while still exhibiting the same structural patterns; languages change over time, but in fairly regular ways. This book focuses on the dynamic processes that create languages and give them their structure and their variance. Joan Bybee outlines a theory of language that directly addresses the nature of grammar, taking into account its variance and gradience, and seeks explanation in terms of the recurrent processes that operate in language use. The evidence is based on the study of large corpora of spoken and written language, and what we know about how languages change, as well as the results of experiments with language users. The result is an integrated theory of language use and language change which has implications for cognitive processing and language evolution.

JOAN BYBEE is Distinguished Professor Emerita in the Department of Linguistics at the University of New Mexico. Her previous publications include *Phonology and Language Use* (Cambridge, 2001) and *Frequency of Use and the Organization of Language* (2007).

Acknowledgements

A researcher's primary debt is to her teachers, students and colleagues, with whom she has exchanged ideas, critiques and inspiration over the years. In my case, a whole community of linguists exploring functional issues, diachronic change, typology and usage effects are cited in the pages of this work for contributing to it through their example and support. This book is based on a forty-year tradition of functional-typological and cognitive linguistics and is dedicated to the men and women who during this period showed the courage to think outside the (black!) box. The book aspires to sum up their work from my own perspective, to apply a consistent set of hypotheses to phonology, morphosyntax and semantics, and to formulate specific new hypotheses about how domain-general processes contribute to the structuring of language.

As for contributions of a more personal nature, I am honoured to name two of my closest friends and colleagues, Sandra Thompson and Rena Torres Cacoullos, who lent personal support and provided their scholarly perspective on drafts of various chapters. I also thank Clay Beckner, who worked with me on Chapters 5 and 8, and came up with much of the data and many of the arguments cited in Chapter 8 (as we collaborated on a paper on a similar topic). I am also grateful to Ben Sienicki for technical support and to Shelece Easterday for constructing the index.

Finally, to my family and friends, who support my activities in many ways, and to my husband, Ira Jaffe, I extend heartfelt thanks.

Contents

Figures

Tables

1   A usage-based perspective on language

'The more it changes, the more it stays the same'

1.1   The nature of language

Sand dunes have apparent regularities of shape and structure, yet they also exhibit considerable variation among individual instances, as well as gradience and change over time. If we want to gain understanding of phenomena that are both structured and variable, it is necessary to look beyond the mutable surface forms to the forces that produce the patterns observed. Language is also a phenomenon that exhibits apparent structure and regularity of patterning while at the same time showing considerable variation at all levels: languages differ from one another while still being patently shaped by the same principles; comparable constructions in different languages serve similar functions and are based on similar principles, yet differ from one another in specifiable ways; utterances within a language differ from one another while still exhibiting the same structural patterns; languages change over time, but in fairly regular ways. Thus it follows that a theory of language could reasonably be focused on the dynamic processes that create languages and give them both their structure and their variance.

A focus on the dynamic processes that create language also allows us to move away from an exclusive focus on linguistic structures and formulate a broader goal: to derive linguistic structure from the application of domain-general processes. In this context, domain-general processes would be those that can be shown to operate in areas of human cognition other than language. The goal of this book is to explore the possibility that the structural phenomena we observe in the grammar of natural languages can be derived from domain-general cognitive processes as they operate in multiple instances of language use. The processes to be considered are called into play in every instance of language use; it is the repetitive use of these processes that has an impact on the cognitive representation of language and thus on language as it is manifested overtly. In this book, then, facts about usage, cognitive processing

1

and language change are united to provide an explanation for the observed properties of linguistic structures.

When linguistic structure is viewed as emergent from the repeated application of underlying processes, rather than given a priori or by design, then language can be seen as a complex adaptive system (Hopper 1987, Larsen-Freeman 1997, Ellis and Larsen-Freeman 2006). The primary reason for viewing language as a complex adaptive system, that is, as being more like sand dunes than like a planned structure, such as a building, is that language exhibits a great deal of variation and gradience. Gradience refers to the fact that many categories of language or grammar are difficult to distinguish, usually because change occurs over time in a gradual way, moving an element along a continuum from one category to another. Continua such as that between derivation and inflection, between function words and affixes, between productive and unproductive constructions, illustrate this gradience. Variation refers to the fact that the units and structures of language exhibit variation in synchronic use, usually along the continuous paths of change that create gradience.

1.2   Gradience and variation in linguistic structure

This section presents some examples of the type of gradience and variation that motivate a view of language as a complex adaptive system. These examples are only a few of the many that one could identify as showing gradience and variation among the members of a particular type of linguistic unit – morphemes (section 1.2.1), language-specific categories – English auxiliaries (section 1.2.2), or variation in instantiations of a particular construction – *I don't* + VERB (section 1.2.3).

1.2.1   *Units: morphemes*

All types of units proposed by linguists show gradience, in the sense that there is a lot of variation within the domain of the unit (different types of words, morphemes, syllables) and difficulty setting the boundaries of the unit. Here I will use morphemes as an example. In their canonical instantiations morphemes involve a constant form associated with a constant meaning. A good example is *happy*, a lexical morpheme. In general, lexical morphemes are less problematic than grammatical morphemes, exhibiting more regularity of form and meaning. However, there are still problematic lexical morphemes which change their meaning and nature depending upon the company they keep. Consider *go*, which often occurs as a simple lexical morpheme, but also occurs in many other constructions, for instance, *go ahead (and), go wrong, go bad, go boom, let's go have lunch,* the famous *be going to* and the quotative *go (and I go 'what do you mean?')* in which its lexical status is quite diminished. We return in Chapter 6 to a discussion of how lexical morphemes become grammatical.

Grammatical morphemes are classically defined as closed class items. Since classes are defined in terms of the properties of constructions, grammatical morphemes are those which are restricted to particular positions in constructions. As a class of unit, grammatical morphemes are highly variable. At the highest level, we find variance across languages in the types of grammatical morphemes that occur. Form and meaning both differ in systematic ways. All languages have function words – non-bound units that express grammatical functions such as tense, aspect, interrogation, negation, and so on. All languages probably also have at least some derivational affixes (Bybee 1985). However, not all languages have inflectional affixes (defined as affixes that belong to obligatory categories). Among those that do have inflection, we traditionally distinguish between agglutinative and fusional languages on the basis of the degree of fusion, allomorphy and irregularity found among the inflectional affixes. Given this range of variation among languages, what similarities do we find among them?

The similarities are apparent in the clines of morphological types, where languages occupy different zones on the cline, ranging from analytic (isolating) to agglutinative to inflectional. The similarities are also apparent in the diachronic processes that create grammatical morphemes, the processes subsumed under the heading 'grammaticalization' (see Chapter 6) by which separate words become affixes and these affixes can become more and more fused with a stem.

Within languages, these same categories can be identified, though rigid distinctions among them are often difficult to make. Gradience is illustrated by the difficulty in determining whether adverbial *–ly* in English is inflectional or derivational (Bybee 1985) or whether the negative particle and its contracted form *–n't* is a clitic or affix (Zwicky and Pullum 1983). Within derivational morphology we find interesting differences not just among affixes, but even considering the same affix in different combinations. The *–ness* suffix in *business* is much less analysable than the same suffix on *happiness*. Hay 2001, 2002 shows that there are even more subtle differences, such as that between the analysability of the suffix in *swiftly* and *softly*.

Grammatical morphemes are bordered by words on the one hand and phonemes on the other. The familiar case of periphrastic expressions using what once were words, such as the perfect *have* + PAST PARTICIPLE, illustrate this gradience, but cases that are not usually cited are cases such as the word *way* in the construction exemplified by *Powell knows how to thread his way between conflicting views.*[1] Since *way* is the only word that can occur in this position in this construction, it qualifies as a grammatical morpheme. However, since it does not fulfil any of the functions traditionally associated with grammatical morphemes, it is more readily recognized as a word. Thus grammatical morphemes that are developing out of words constitute one side of the gradient, and

on the other side are grammatical morphemes that are losing their meaningful status and becoming just part of the phonology of the word. Hopper 1994 discusses a number of such cases (see also Greenberg 1978b). One example is the second syllable of English *seldom*, which was previously the dative plural marker attached to the adjective *seld* 'strange, rare' and now is a meaningless part of the word.

The variation and gradience in the category of 'grammatical morpheme' is a direct result of the processes of change that affect morphemes and shape their properties of form and meaning. Lexical morphemes can become grammatical morphemes in the process of grammaticalization (as when the lexical morpheme *go* becomes part of the future construction *be going to*), and in this process gradually become more dependent upon and eventually fused with surrounding material. Derivational morphemes are formed when two words are used together in a compound-like fashion; thus –*ly* came from the noun *lič-* meaning 'body' which produced a compound meaning 'having the body of'. The second part of this compound gradually reduced, spread to more and more nouns and adjectives and generalized its meaning in the process.

Of course, these processes of change are well known and quite taken for granted. What is not so well appreciated, however, is what they tell us about the cognitive processing that is used in language. They tell us something about how language use affects storage in memory and the organization of that storage (Bybee and McClelland 2005, McClelland and Bybee 2007). In the chapters of this book we will be examining the tendencies that are at work when language is being processed. Rather than taking the gradience just illustrated as a descriptive problem, let us consider it the very essence of the phenomenon and think of language as ever being affected by language use and the impact that experience has on the cognitive system.

1.2.2   Language-specific categories that are heterogeneous and gradient: the English auxiliary

The English auxiliary sequence is worthy of close scrutiny because it appears to be a very good example of a well-behaved linguistic structure that is involved in certain clear rules. In Chapter 7 I examine the way this structure and the associated rules or constructions (of subject–auxiliary inversion and negation) came into being in the sixteenth century. There we will see that a number of gradual changes, some of them only remotely related at first, led to the formation of the auxiliary and its related constructions. This study brings to light the fact that the element that inverts with the subject and takes negation following it is actually a diverse structural class, including the set of modal auxiliaries (more comments on this set below) which appear with an unmarked main verb form; two constructions that each take a different form of the main verb: the

Progressive (BE + ING) and the Perfect (HAVE + EN); and one element, the copula, which is actually the main verb in the predicates in which it appears. This category also formerly contained (and still does in some dialects) the possessive *have* and indeed many other verbs. Thus the members of the category of auxiliary items in English are quite diverse, being neither structurally nor functionally uniform.

Moreover, the category itself has less than discrete boundaries. The elements mentioned – the modals, the Progressive *be*, the Perfect *have* and the copula – are quite categorically members of this class of items, but the verbs *dare* and *need* sometimes behave as though they were members of the category and sometimes behave as though they were ordinary main verbs. This gradience is not just some passing stage; rather these two verbs have straddled the two categories since the time the category of auxiliary started to differentiate from the category of main verb (some five centuries ago, see Chapter 7).

In addition, the members of the category of modal auxiliary are also diverse and show variation, especially in function. While most express modality, either agent-oriented, ability or root possibility (*can* and *could*), obligation (*must, shall* and *should*) or epistemic (*may, might, could*), some also express tense (*will* and *shall* for future) or tense and aspect (*would* for past habitual).

This class of items with very similar structural properties expresses a wide range of different meanings. Such a category is not unusual in the languages of the world. Bybee 1986 surveyed tense, aspect and modality inflections in fifty languages and found that it is actually rather uncommon for position classes to correspond directly to meaning categories. This heterogeneity is not specific to affixes and auxiliaries; prepositions also show many differences in behaviour, with *of*, the most common, often not behaving much like a preposition at all (Sinclair 1991) and complex prepositions (such as *on top of, in spite of*) showing mixed behaviour between containing two prepositions and a noun and functioning as a unit (Chapter 8).

1.2.3   Specific instances of construction vary:
I don't know, I don't inhale

The types of gradience and variation discussed in the preceding subsections are well-known from the literature (as noted above), but the final type of gradience I want to discuss has only more recently received attention as a phenomenon that a linguistic theory needs to reckon with. In this subsection we focus on the fact that, at times, specific instances of constructions (with particular lexical items included in them) take on behaviour different from the general construction.

Consider the two expressions *I don't know* and *I don't inhale*. They appear to be structurally identical, each one having a first-person-singular pronoun

followed by the negated form of the *do* auxiliary and an unmarked main verb. They both exhibit the same phonetic variation in that the initial stop of *don't* can become a flap and the final [t] is usually deleted in both cases. But in addition to this, the first expression *I don't know* also has number of other variant properties that the second expression does not share. Even though *I don't know* can certainly have the meaning that is predictable from the sum of its parts, it is also often used as a discourse marker, mollifying the force of the previous assertions and letting the listener know that the speaker is willing to give up the floor (Scheibman 2000). In this discourse-pragmatic usage, the phrase is also more likely to have further phonetic reduction than in its more semantically transparent usage. The further reduction involves the vowel of *don't* which becomes a schwa. The most extreme reduction which occurs in this phrase is the loss of the initial stop [d]. Neither of these changes occur when the main verb is a less frequent verb such as *inhale* (Bybee and Scheibman 1999). See Chapter 2 for further discussion of reduction and semantic changes in high-frequency expressions.

1.2.4   *The role of gradience and variation*

To these few examples, one could add many more: the difficulty of defining units such as 'segment', 'syllable' and even 'word', the problem with the notion of 'clause' when clauses take so many shapes, and the fact that grammaticality judgements show gradience and variation across speakers. The existence of gradience and variation does not negate the regular patterning within languages or the patterning across languages. However, it is important not to view the regularities as primary and the gradience and variation as secondary; rather the same factors operate to produce both regular patterns and the deviations. If language were a fixed mental structure, it would perhaps have discrete categories; but since it is a mental structure that is in constant use and filtered through processing activities that change it, there is variation and gradation.

1.3   Domain-general processes

Language is one of the most systematic and complex forms of human behaviour. As such it has given rise to many different theories about what it is used for (thinking vs. communicating), how it has evolved (abruptly or gradually), where its structure comes from (innate structures vs. language use) and what types of processes underlie its structure (those specific to language vs. those applicable in many cognitive domains). Here we consider the last question – are the processes that give us linguistic structure specific to language or are they processes that also apply in other cognitive domains? The best strategy for answering this question is to start first with domain-general processes and see

how much of linguistic structure can be explained without postulating processes specific to language. If this quest is even partially successful, we will have narrowed down the possible processes that have to be specific to language. The opposite strategy of assuming processes specific to language will not lead to the discovery of how domain-general processes contribute to linguistic structure.

As mentioned above, a consequence of viewing language as a complex adaptive system and linguistic structure as emergent (Lindblom et al. 1984, Hopper 1987) is that it focuses our attention not so much on linguistic structure itself, as on the processes that create it (Verhagen 2002). By searching for domain-general processes, we not only narrow the search for processes specific to language, but we also situate language within the larger context of human behaviour.

The domain-general cognitive processes studied in this book are categorization, chunking, rich memory storage, analogy and cross-modal association. This list is not meant to exhaust the cognitive processes involved in language, nor to deny that there might be processes specific to language that will be discovered; the list represents the processes that have proven useful in understanding some aspects of language that have particularly interested me.

Categorization is the most pervasive of these processes as it interacts with the others. By categorization I mean the similarity or identity matching that occurs when words and phrases and their component parts are recognized and matched to stored representations. The resulting categories are the foundation of the linguistic system, whether they are sound units, morphemes, words, phrases or constructions (see Chapters 2, 4, 5 and 8). Categorization is domain-general in the sense that perceptual categories of various sorts are created from experience independently of language.

Chunking is the process by which sequences of units that are used together cohere to form more complex units. As a domain-general process chunking helps to explain why people get better at cognitive and neuromotor tasks with practice. In language, chunking is basic to the formation of sequential units expressed as constructions, constituents and formulaic expressions. Repeated sequences of words (or morphemes) are packaged together in cognition so that the sequence can be accessed as a single unit. It is the interaction of chunking with categorization that gives conventional sequences varying degrees of analysability and compositionality (Chapters 3 and 8).

Rich memory refers to the memory storage of the details of experience with language, including phonetic detail for words and phrases, contexts of use, meanings and inferences associated with utterances. Categorization is the process by which these rich memories are mapped onto existing representations (Chapter 2). Memory for linguistic forms is represented in exemplars, which are built up from tokens of language experience that are deemed to be identical. The primary claim of exemplar representation is that each experience with language has an impact

on cognitive representations. Non-linguistic memories also have an impact on cognitive representations and on neurological structure (Nader et al. 2000).

Analogy is the process by which novel utterances are created based on previously experienced utterances. Analogy also requires categorization; the parts of previously experienced tokens must be parsed into units that are aligned and categorized before novel utterances can be formed from them. Analogy is domain-general and has been studied in terms of relational structures on visual stimuli, such as scenes, shapes and colours (Gentner 1983, Gentner and Markman 1997).

The list of domain-general processes also includes the ability to make cross-modal associations that provide the link between meaning and form. Ellis (1996) discusses this most basic principle as James' Law of Contiguity (James 1950 [1890]) by which co-occurring experiences tend to be associated in cognition. Ellis goes on to point out that

The implicit, automatic pattern-detection processes that occur within these modalities of representation entail that any such cross-modal association typically occur between the highest chunked levels of activated nodes. Thus to extend Morton's (1967) example, the adult looking at his or her watch when the post falls through the mail slot each morning learns an association that mail time is 8:30 a.m., not one between envelopes and the big hand of the watch. (1996: 110)

Thus meaning is assigned to the largest chunk available – a word, a phrase or a construction. Note that inferences made from the context of particular utterances can also come to be associated with particular sequences, giving rise to changes in meaning (see Chapters 3, 6, 8 and 10).

Chapters 2 through 5 of this book discuss these domain-general processes and the way that their iterative application in language use creates the categories and units of language, sequential structures such as constructions and constituents. It is also shown that variations in analysability and compositionality as well as the productive and creative use of language are derivable from these same processes. Chapters 6 through 8 examine in more detail how these same processes apply in cases of language change, especially in cases of grammaticalization, in the creation of new constructions and in changes in constituent structure. Chapter 10 is devoted to discussing the consequences of these proposals for our understanding of the meaning of grammatical categories. Chapter 11 considers the way similarities among languages arise through application and interaction of domain-general processes during language use in particular cultural contexts.

1.4   Usage-based grammar

In Bybee 2006a I proposed that grammar be thought of as the cognitive organization of one's experience with language. To cast this in terms that linguists