

REVIEW ARTICLES

Language, usage and cognition. By JOAN BYBEE. Cambridge: Cambridge University Press, 2010. Pp. ix, 262. ISBN 9780521616836. \$39.99.

Reviewed by HOLGER DIESEL, *Friedrich-Schiller-Universität Jena**

1. INTRODUCTION. There is a long tradition in linguistics and philosophy of analyzing language without reference to usage and experience. This tradition is reflected in Chomsky's famous division between competence and performance and Saussure's related distinction between *langue* and *parole*, which have influenced linguistic research for many decades. This view of language, however, has been called into question by usage-based linguists who have emphasized the importance of communication, cognition, and processing for the development and organization of grammar. In the usage-based approach, grammar is seen as an emergent system consisting of fluid categories and dynamic constraints that are in principle always changing under the influence of general cognitive and communicative pressures of language use.

Joan Bybee is one of the pioneers of the usage-based approach, which has its roots in functional and cognitive linguistics (e.g. Givón 1979, Hopper 1987, Langacker 1987) and is related to various strands of research in cognitive science (e.g. Bates & MacWhinney 1989, Bod 2003, Clark 1996, Elman et al. 1996, Tomasello 2003). Thus far, B's work has been mainly concerned with morphology (e.g. Bybee 1985) and phonology (e.g. Bybee 2001); but in the current book the focus is on larger grammatical patterns (i.e. linguistic units that exceed the single lexeme). The book provides an overview of some of B's earlier research on grammar and grammaticalization and presents some new corpus studies and usage-based analyses from the current literature.

B's view of grammar rests on central assumptions of construction grammar, which has made important contributions to the usage-based approach (e.g. Goldberg 2006, Langacker 2008). In fact, usage-based linguists have drawn so frequently on concepts of construction grammar that the two approaches are often presented as a unified theory (e.g. Dabrowska 2004, Diessel 2004, Tomasello 2003).¹ In this approach, grammar consists of emergent form-function units, or 'fluid constructions' (Goldberg 2006), which are related to each other by probabilistic links that are determined by their similarity and cooccurrence in usage. Since constructions involve the same cross-modal associations of form and meaning as words and morphemes, they are subject to the same cognitive processes as lexical expressions. Thus, many of the usage-based principles that B analyzed in her previous research on morphology and phonology can also be applied to the analysis of larger grammatical patterns, that is, constructions.

The general goal of the book is 'to derive linguistic structure from the application of domain-general processes' (1), or, as Lindblom and colleagues (1984:187) put it, 'DE-

* I would like to thank Volker Gast, Florian Haas, Katja Hetterle, Martin Schäfer, Karsten Schmidtke-Bode, and Daniel Wiechmann for helpful comments and suggestions.

¹ Note, however, that construction grammar subsumes a whole family of grammatical theories that are not generally compatible with the usage-based approach. In fact, one of the earliest and most influential construction-based theories, that is, Fillmore and Kay's version of construction grammar, explicitly accepts the division between competence and performance and thus does not share the dynamic view of grammar that underlies the usage-based approach (cf. Fillmore & Kay 1993). However, other varieties of construction grammar, notably Langacker's cognitive grammar and Goldberg's construction-based theory of verb-argument structure, have taken a usage-based perspective and have made significant contributions to this approach.

RIVE LANGUAGE FROM NONLANGUAGE!'. Domain-general processes are cognitive mechanisms that are relevant not only for language but also for other cognitive phenomena such as vision and thought. Drawing on general research in cognitive science, B relates the analysis of language use and change to general processes of cognition.

The discussion in the book is divided into eleven chapters. The introductory chapter presents some central assumptions about the usage-based approach and introduces the domain-general cognitive processes B considers important for the emergence of linguistic structure. The four following chapters are concerned with four of the five cognitive processes presented in the introduction: rich memory (Ch. 2), chunking (Ch. 3), analogy (Ch. 4), and categorization (Ch. 5); but from Ch. 6 on, the structure of the book is difficult to grasp. Some chapters are concerned with particular mechanisms of change (e.g. reanalysis; Ch. 7), others present case studies on English auxiliaries and modals (Chs. 7 and 9), and yet others cover an array of topics that are only loosely related (e.g. Chs. 6 and 11). Since many issues are discussed across several chapters, I summarize the main points of the book in ten theses on the usage-based approach that underlie the entire discussion.

2. TEN THESES ON THE USAGE-BASED APPROACH.

Thesis 1: Language is grounded in domain-general cognitive processes. This is the most general thesis, providing a foundation for the whole approach. B conceives of language as a 'complex adaptive system' that is shaped by 'domain-general cognitive processes' involved in usage. The same idea has been expressed in related research in linguistics and cognitive science (e.g. Bates & MacWhinney 1989, Elman et al. 1996, Lindblom et al. 1984, Tomasello 2003) and can be seen as a general feature of the usage-based approach, which contrasts sharply with the generativist view of a particular language faculty consisting of language-specific categories and constraints (for a recent discussion of the language faculty, see Pinker & Jackendoff 2005). In the usage-based approach, there is no (innate) language faculty, and grammar is grounded in general cognitive processes that are involved not only in language but also in nonlinguistic cognitive activities (e.g. visual and auditory perception, nonlinguistic memory, joint attention, and reasoning).

Elaborating this general framework, B proposes five cognitive mechanisms (or 'cognitive processes') that influence the use and development of linguistic structure in specific ways: (i) categorization, which is the most general cognitive mechanism that accounts for the identification of individual tokens through comparison with previously established categories; (ii) chunking, which refers to the formation of sequential units through repetition or practice; (iii) rich memory, which refers to the storage of detailed information from experience; (iv) analogy, which characterizes the mapping of an existing structural pattern onto a novel instance; and (v) cross-modal association, which refers to the cognitive capacity to link form and meaning.

B's proposal to derive language from general cognitive processes seeks to ground the usage-based theory of language in a general theory of cognition. One principal problem of current research in functional and cognitive linguistics is that explanations of linguistic structure often refer to very different aspects of usage and cognition. Some studies refer to discourse factors such as coherence and grounding (e.g. Givón 1995, Hopper & Thompson 1984), others refer to semantic and conceptual factors such as iconicity and construal (e.g. Haiman 1985, Langacker 2008), and yet others allude to processing factors such as automatization and efficiency (e.g. Bybee 2002, Hawkins 2004). There are so many functional and cognitive motivations that usage-based expla-

nations of linguistic structure have been rightfully criticized for being arbitrary and ad hoc (see Newmeyer 1998). Obviously, what is needed is a more systematic approach that differentiates between different levels of analysis and thereby constrains the explanatory power of the usage-based theory of language.

B's proposal to analyze language in terms of specific cognitive processes can be seen as a first step in this direction. It establishes a framework for the analysis of linguistic structure that rests on general research in cognitive science. Categorization, chunking, rich memory, analogy, and cross-modal association provide a cognitive foundation for the analysis of more specifically linguistic activities such as reference tracking, discourse management, sentence processing, and the encoding of meaning. The five cognitive processes, however, are not at the same level. Categorization is an overarching mechanism that is involved in almost all cognitive activities in language; it interacts with chunking, analogy, and cross-modal association and underlies B's view of rich memory, which she basically defines in terms of exemplar-based categorization. Moreover, B points out that the cognitive processes presented in the introduction are not meant to be exhaustive; there may be other cognitive processes that are relevant for the usage-based analysis of language. In fact, one aspect that seems to be missing from her list (though it is implicit in the discussion) is the ability to understand other people's intentions and perspectives. There is evidence that 'perspective taking' and 'mind reading' are fundamental capacities of the human mind that influence the use and development of language in specific ways (cf. Clark 1996, Sperber & Wilson 1995, Tomasello 2003). Although the details of B's proposal are debatable, the general framework outlined in this book provides a good starting point for the development of a more systematic and comprehensive theory of the usage-based approach that is grounded in general mechanisms of the mind.

Thesis 2: The synchronic analysis of grammar cannot be separated from the analysis of diachronic change. Since the beginning of modern linguistics (i.e. since Saussure's *Cours de linguistique générale*), synchrony and diachrony have been commonly conceived of as two separate subfields of linguistics addressing different research questions and employing different methodologies. In the usage-based approach, however, linguistic structure is generally seen in the light of its development. Synchrony and diachrony are reunited as they were in the prestructuralist period of linguistics (e.g. Paul 1880).

In accordance with this view, the current book emphasizes the importance of diachronic processes for the analysis of grammar. It analyzes grammatical categories and constructions in terms of their development (Chs. 4 and 6), it offers a diachronic analysis of constituency (Chs. 3 and 8), and it provides a historical explanation for language universals (Ch. 11). Since the usage-based model of grammar is a dynamic system with emergent categories and constraints, usage-based linguists naturally focus their attention on the processes that create grammar in the course of language change. This is why grammaticalization is so important for the usage-based approach. According to B, grammaticalization has 'demystified' the mental view of grammar that underlies the generative approach and has provided strong evidence for the assumption that grammatical patterns are shaped by general cognitive processes (especially Ch. 6). Moreover, grammaticalization helps to explain why (many) syntactic patterns are lexically specific (thesis 5) and why grammatical categories are gradient (thesis 6).

Some generativist linguists have criticized the usage-based analysis of (synchronic) grammar in terms of diachronic change because in their view the core elements of grammar can only change in language acquisition, that is, when children acquire the

principal parameters of their native language (see e.g. Newmeyer 1998). In the usage-based approach, however, the source of diachronic change is (adult) language use rather than L1 acquisition (see e.g. Bybee 2001, Croft 2000). In support of this hypothesis B presents three arguments (114–19): first, she points out that children's deviant uses do not seem to persist into adulthood; second, she presents some evidence for change in adult language; and third, she notes that children's errors are distinct from changes in historical development, making a direct causal link between ontogeny and diachrony unlikely. It seems that language development results from small changes in the speech of adult speakers that gradually replace older patterns through lexical and geographical diffusion (Ch. 6).

Thesis 3: Frequency is an important determinant of language change. Like B's earlier work on morphology and phonology, the current book emphasizes the importance of frequency for cognition and language (cf. Bybee 2007, Bybee & Hopper 2001). Of course, frequency itself is not a cognitive phenomenon; the term simply denotes the occurrence of objects or events in a particular domain or time frame. But frequency of occurrence is an important factor in almost all cognitive processes that are involved in usage and development: it underlies the emergence of exemplar-based categories (Ch. 2); it influences analogy and pragmatic inference (Chs. 4 and 10); and it has a major impact on grammaticalization and other aspects of diachronic change (especially Chs. 3 and 6) (see also Diessel 2007 for a recent review of frequency effects in grammar, usage, acquisition, and change).

Building on her earlier research in morphology, B shows that frequency can influence the development of larger grammatical units in various ways. Two frequency effects are in the focus of this book: (i) the reduction effect, which refers to the articulatory reduction and coarticulation of frequent expressions and constructions (Ch. 3), and (ii) the preserving effect, which refers to the survival of frequent linguistic patterns in analogical change (Ch. 4).

The reduction effect of token frequency seems to have two major psychological sources (38–43): on the one hand, speakers reduce frequent words and word sequences simply because frequent expressions are more strongly represented in memory and therefore more easily accessible in speech production; and on the other hand, speakers seem to produce frequent expressions with less articulatory effort because they know that frequent expressions are more easily predictable for the addressee. The first aspect is a purely speaker-oriented process that is determined by repetition and practice, whereas the second aspect is a listener-oriented process that involves a certain amount of mind reading (see Bell et al. 2009 for a recent study examining these issues).

The preserving effect of token frequency results from the interaction between the strength of linguistic representations in memory and the power of analogy. As B demonstrates with historical data from the development of questions and negative sentences in Early Modern English, frequent grammatical patterns are often exempt from analogical leveling because they are so deeply entrenched in memory that they do not change (Chs. 4 and 7). As a result, grammars often include archaic structural patterns that have resisted the pressure from analogical change and are synchronically analyzed as 'exceptions' to the general rule, that is, the new grammatical pattern that was established by analogy (e.g. the occurrence of postverbal negation in Modern English after the emergence of preverbal negation with *do* in Early Modern English, for example, *They don't know what they do* vs. *They know not what they do*).

Although B emphasizes the importance of token frequency for the emergence of grammar, the current book makes it clear that frequency alone is not sufficient to ex-

plain the dynamic processes of usage and diachronic change. A main feature of the current book is that it considers the role of frequency in conjunction with other factors that influence the use and development of linguistic structure.

Thesis 4: Linguistic categories are based on concrete tokens. An important aspect of token frequency is its role in categorization. Consistent with B's earlier work on categorization in phonology (e.g. Bybee 2001), the current book argues that grammatical categories and constructions are based on concrete tokens or exemplars (especially Ch. 2). In such an exemplar-based view of categorization, every linguistic expression in usage contributes to the representation and development of linguistic categories and constructions. The theory rests on the assumption that language users are endowed with a very rich memory system that allows them to store large amounts of information, which may even survive (in memory) if this information is subsumed under a generalization. As a consequence, novel tokens are not generally categorized by abstract schemas (or rules), but are often licensed by individual tokens or token clusters that may be stored in memory in addition to high-level generalizations.

This view of linguistic categorization is fundamentally distinct from the traditional assumption that linguistic categories are highly abstract entities that are applied to every novel utterance. In the generativist view of linguistic categories, concrete tokens are irrelevant for linguistic categorization if they are subsumed under a generalization. The abstractness of linguistic categories is, at least in part, motivated by the assumptions that the language system is most efficient if it rests on maximally abstract representations so that the amount of information in memory can be reduced to the absolute minimum.

B takes the opposite view. On her account, memory is cheap and computation is costly. Therefore, language users will often draw on concrete tokens and low-level generalizations to license a particular structure. This does not mean that abstract categories are irrelevant for linguistic categorization. There is no doubt that grammar includes highly abstract representations; but they may be less important for the use of novel utterances than commonly assumed in linguistics. In fact, B argues that not every descriptive generalization unraveled by some clever linguist is psychologically real in the sense that it represents the language users' underlying linguistic knowledge (Ch. 5). Grammatical constructions may share a particular morphosyntactic property, such as a particular word order or case marker, not because they are licensed by the same schemas (or rules), but because they are diachronically related. As B points out, 'since new constructions develop out of existing constructions, the properties of existing constructions are carried over into new ones over time' (102). Therefore, historically related constructions can be structurally similar without being licensed by a (synchronic) generalization. On this account, the scientific discovery of a descriptive generalization is not sufficient to posit the existence of a 'mental rule'; rather, B suggests that observational generalizations have to be tested in order to find out if they are psychologically real (102–3).

Thesis 5: Syntactic structure is lexically specific. Since grammar emerges from the language users' experience with concrete tokens, that is, concrete words and utterances, syntactic structure is commonly tied to specific lexical expressions. This is another fundamental difference between the generativist model of grammar and the usage-based approach. In the generativist theory of grammar, syntactic structure is abstract and independent of concrete lexical expressions; grammar and lexicon are strictly separated in this approach. But in the usage-based model of grammar, syntactic structure is lexically specific: most grammatical constructions are associated with specific lexical

items, which is of course a consequence of the fact that grammar emerges from our experience with concrete tokens. Since grammatical categories and constructions emerge from concrete exemplars, they are tied later to particular lexical expressions that are (frequently) experienced in particular positions.

Drawing on recent research from corpus linguistics, B shows that the ‘slots’ of verb-argument constructions are often biased toward particular expressions. For instance, the English resultative construction (e.g. *John pushed the door open*) subsumes various lexical patterns that are organized around specific words (Boas 2003). In one of these patterns, the verb *drive* occurs with an animate object that is commonly expressed by a personal pronoun (e.g. *me, you, him, her*) followed by the adjective *crazy* or some semantically related expression such as *mad* or *up the wall*. Following Boas (2003), B argues that the resultative construction is associated with specific words that cannot be excluded from the (grammatical) representation of this construction.

Lexical expressions that are associated with a particular slot in a construction influence the categorization of novel expressions in this position and the development of constructions in diachronic change. As B and others have amply demonstrated, grammaticalization is an item-specific process that involves the reanalysis of lexically specific structures, that is, constructions including a particular lexical item that is reanalyzed as a grammatical marker (see e.g. Bybee 2003). The same holds true for analogical change, which commonly proceeds in an item-specific fashion, referred to in historical linguistics as lexical diffusion (Tottie 1991). For instance, the current book shows that the emergence of a new negative pattern in Early Modern English (e.g. *NP ne Verb ...* → *NP does not Verb ...*) affected different verbs at different times until all verbs were subsumed under the new pattern except for the most frequent ones, that is, the present day auxiliaries and modals and a few lexicalized forms that have preserved the old pattern of postverbal negation (e.g. *can not, have not, is not, know not, make no mistake*) (69–71, 123–24).

The lexical-specific nature of grammar is perhaps most obvious in early child language. As Tomasello, Lieven, and colleagues have demonstrated through both experimental and observational studies, children’s early grammatical constructions are organized around particular lexemes that are associated with an open slot to form multiword expressions (Tomasello 2003, Lieven et al. 1997). B reviews some of this research, notably the research by Goldberg and colleagues, arguing that the acquisition of grammatical patterns is facilitated if children can associate a particular construction with one or more high-frequency expressions (91). There is also evidence from sentence processing (MacDonald et al. 1994) and syntactic priming (Pickering & Branigan 1999) that indicates the lexical-specific nature of syntax; but this research is not considered in B’s book.

Thesis 6: Grammatical categories are gradient. Since linguistic categories are derived from concrete utterances, they usually overlap with neighboring categories derived from similar tokens; that is, there are no clear-cut boundaries between them. The structure of linguistic categories is one of the most hotly debated issues in linguistics (e.g. Lakoff 1987). When Eleanor Rosch and other cognitive psychologists established prototype theory as an alternative to the classical approach to categorization, cognitive and functional linguists adopted the notion of prototype for the analysis of linguistic categories, including categories of morphosyntax. In accordance with this approach, B emphasizes that linguistic categories are gradient and organized around prototypical members (Chs. 1 and 5).

More recent work on categorization has been concerned with the relationship between prototype theory and the exemplar model (see Murphy 2002 for a review). The two theories are now commonly seen as complements (rather than alternatives) that focus on different aspects of categorization: exemplar theory emphasizes the importance of experience and individual tokens for categorization, whereas prototype theory is concerned with abstract summary representations that are derived from concrete tokens but can license categorization processes independently of them (Abbot-Smith & Tomasello 2006). B does not explicitly consider the relationship between prototype theory and the exemplar model, but her discussion suggests that linguistic categorization involves both concrete tokens and abstract summary representations that are permanently stored in memory as generalizations over individual tokens.

Although B seems to acknowledge the role of abstract representations for linguistic categorization, she emphasizes the importance of exemplar theory for the analysis of linguistic structure. As pointed out above, exemplar theory explains why language abounds with prefabricated chunks and collocations not licensed by a schema (i.e. an abstract summary representation), and it provides a straightforward account for the emergence and gradience of linguistic categories. If grammar is shaped by our experience with concrete tokens, that is, concrete words and utterances, then linguistic categories must have a gradient structure, which commonly reflects the course of their development. As B notes, grammaticalization usually consists of a sequence of many small changes whereby an existing construction is gradually transformed into a new one (especially Ch. 6). Since the source construction and other intermediate constructions are not generally discarded in this process, they often constitute a chain of related grammatical patterns with fuzzy boundaries between them.

Thesis 7: Constituent structure is determined by chunking. That linguistic categories are gradient has become a standard assumption of research in grammaticalization and cognitive linguistics. There is one aspect of grammatical structure, however, that has received little attention in this context: constituency. While grammatical word classes and grammatical relations are commonly seen as gradient categories, constituent structure is still often analyzed in terms of classical categories represented in phrase structure trees or other discrete representations (e.g. boxes in head-driven phrase structure grammar (HPSG)).

In a series of recent papers, however, B has argued that constituent structure exhibits the same gradience and indeterminacy as other aspects of grammar (Bybee & Scheibman 1999, Bybee 2002, Beckner & Bybee 2009). Referring to this research, the current book argues that the hierarchical organization of phrase structure is determined by chunking—a cognitive mechanism that transforms a sequence of separate entities, for example, a sequence of words, into a holistic unit through repetition (Chs. 3 and 8).

Of course, syntactic constituents have a semantic basis (Ch. 8). It is well known that the linear arrangement of words and phrases is motivated by their meaning: linguistic expressions that are semantically related tend to occur adjacent to each other (Behaghel 1932). But B points out that the semantic basis of constituency is not sufficient to explain the hierarchical organization of syntax. In her analysis, constituents are processing units, which have emerged from recurrent strings of linguistic elements that are bound together through repetition or chunking.

Since the formation of chunks is a continuous process, the emerging phrases exhibit varying degrees of cohesion. Other things being equal, smaller chunks (e.g. *the dog*) tend to be more tightly organized than larger ones (e.g. *the old dog over there that is barking*) because they are more frequent, suggesting that constituency is a gradient con-

cept just like any other grammatical category. Since the emerging chunks (or phrases) can be included in each other, the combination of smaller and larger chunks creates a hierarchical structure that (structuralist) linguists represent in phrase structure trees (35).

B emphasizes the importance of chunking for the emergence of hierarchical structure, but she also notes that chunking can eliminate or reduce an existing phrase structure configuration (Ch. 3). More specifically, she argues that chunks of concrete lexical expressions may lose the connection to regular phrase structure patterns if they are very frequent (44–45). For instance, as shown in Bybee & Scheibman 1999, an important study that is summarized in Ch. 3, the frequent occurrence of *I don't know* [CLAUSE] and *Why don't you* [CLAUSE] in complex sentences has altered their structure and meaning. Comparing these expressions to other instances of the same type (e.g. *Peter didn't say* [CLAUSE], *Why did John* [CLAUSE]), Bybee and Scheibman found that *I don't know* and *Why don't you* are phonetically more reduced and semantically less transparent than other instances of the same construction, suggesting that these expressions have become 'autonomous' in the sense that they have lost (or weakened) their connections to the regular phrase structure patterns from which they derive (41–44).

Following Langacker (1987), B argues that the autonomy of lexical chunks can be measured by two aspects: (i) (semantic) compositionality, that is, the degree to which the meaning of a chunk is derived from the meaning of its components, and (ii) (structural) analyzability, that is, the degree to which language users recognize the contribution of individual lexemes. In the current example, both aspects are affected by chunking. In their phonetically reduced forms, *I don't know* and *Why don't you* are no longer used to express literal negation, serving instead particular pragmatic functions as epistemic markers and markers of the illocutionary force, which cannot be directly inferred from their lexical components.

Thesis 8: The meaning of grammatical markers and constructions is polysemous. The bulk of B's book is concerned with the analysis of linguistic structure and considers the semantic side of grammar only in passing; however, there is one chapter that is devoted to the meaning of grammatical markers and constructions (Ch. 10). In this chapter, B outlines a usage-based analysis of 'grammatical meaning' in contrast to the structuralist analysis of meaning as developed by Roman Jakobson. In the structuralist approach, the meaning of grammatical categories, such as aspect or number, is analyzed in terms of binary oppositions that are defined by a limited number of semantic features (e.g. perfective vs. imperfective, singular vs. plural). There is a tendency to characterize the meaning of each category member by an invariant set of semantic features across different contexts of use.

Challenging this approach, B argues that grammatical meaning does not involve binary semantic oppositions, as proposed by Jakobson, but reflects the language users' experience with particular situations. Since our experience of the world is open-ended, the meaning of linguistic expressions cannot be adequately analyzed by means of a restricted set of semantic features; rather, what is needed is a dynamic theory of meaning, in which the semantic features of linguistic expressions are determined by their use in different situations and contexts. Since linguistic expressions are never tied to one particular situation—that is, they are always used in multiple situations and contexts—they are usually polysemous.

There is a large body of research on the multifunctionality and polysemy of linguistic expressions in cognitive linguistics (cf. Lakoff 1987, Langacker 1987) providing the (implicit) background for B's discussion in this chapter. What B emphasizes in her analysis of grammatical meaning is that the semantic features of grammatical markers are cru-

cially determined by constructions: 'Grammatical morphemes are always part of a construction and their meaning can only be understood as deriving from the meaning of the whole construction' (176). A good example is provided by the development of future-tense auxiliaries from verbs of volition, obligation, and motion, which commonly originates in constructions with an implicit future-time meaning that is later absorbed by the tense marker (172–73). Since the original meaning is commonly retained in certain contexts, however, both constructions and grammatical markers tend to have multiple meanings, which together constitute a semantic network of grammatical markers and constructions. This network is constantly modified and restructured through small changes that are caused by the language users' ever ongoing experience with language.

Thesis 9: Linguistic productivity involves analogy rather than rules. If linguistic categories are gradient and lexically specific, linguistic productivity cannot be analyzed in terms of traditional rules. In generative linguistics, grammar is commonly characterized as a closed, deductive system that has been analyzed in terms of formal language theory (cf. Chomsky 1957). On this view, grammar involves discrete symbols and categorical rules that allow for no exceptions.

In the usage-based analysis of linguistic productivity, traditional rules have been abandoned and replaced by schemas (Bybee 1995, Langacker 2008). A schema is a grammatical template, or abstract construction, that has evolved through generalization over concrete tokens. Since there are usually multiple schemas that a speaker can potentially use to produce (or comprehend) a novel utterance in a particular situation, linguistic productivity can be defined as the process whereby language users select a specific schema from a set of alternatives (Langacker 2008). The choice (or selection) of a schema is determined by a cognitive process that B and others have characterized by the notion of analogy, which in turn is influenced by two factors: (i) type frequency and (ii) similarity (Ch. 4).

B (1985) was one of the first to argue that linguistic productivity is influenced by type frequency (which is now a common assumption among usage-based linguists). The larger the number of types subsumed by a particular schema, the greater the likelihood that it will be selected to license a novel expression or construction. For instance, the regular past tense of English subsumes a very large number of verb types, which is one reason why the *-ed* past tense is more productive than the irregular past-tense forms (cf. Bybee 1995). The same analysis holds for constructions. As argued in Ch. 4, the more types of lexical expressions are licensed by the slots of a particular construction, the more productive is its (future) use (57). For instance, *if*-conditional clauses (e.g. *If I had been there, ...*) are more productive than verb-first conditional clauses (e.g. *Had I been there, ...*) because the latter can only be formed with a few modals.

The second factor that influences the productivity of a schema is similarity, which in B's earlier work on the English past tense was defined in phonological terms: the regular past tense is an 'open schema' that does not have any phonological constraints, but the irregular past-tense forms are based on phonological templates (e.g. *sing/sung*, *ring/rung*, *drink/drank*) that influence their productivity. The more phonological features a (novel) verb shares with one of the existing templates of irregular verbs, the greater the likelihood that it will be inflected according to this pattern (e.g. *bing/bung*) (see Bybee & Moder 1983). In the current book, B emphasizes the importance of semantic similarity between lexical expressions for the productivity of a particular slot in a construction. For instance, as pointed out above, the final slot of the *drive-somebody-crazy* construction is restricted to expressions that are semantically similar to 'crazy' (cf. *to drive somebody insane* vs. *to drive somebody sane*). Analyzing parallel construc-

tions in Spanish, B concludes that analogy in grammar is often lexically specific and constrained by the semantic features of constructions.

At the end of Ch. 4, B summarizes the main differences between analogy and (symbolic) rules: (i) analogy is crucially influenced by (type and token) frequency, whereas rules are not affected by frequency; (ii) analogy is influenced by similarities between lexical expressions, whereas rules are insensitive toward specific lexical items; (iii) analogy can involve both form and meaning, whereas syntactic rules are usually restricted to structure; and (iv) analogy gives rise to a gradual notion of productivity, whereas rule-based productivity is absolute (73–74).

Thesis 10: Language universals are dynamic and statistical rather than static and absolute. The final chapter of B's book addresses the issue of language universals. In the generative approach, language universals are determined by aspects of the language faculty and are therefore absolute. As linguistic typologists have repeatedly noted, however, absolute language universals are either trivial ('all languages have words') or difficult to verify (see e.g. Dryer 1997). In fact, many language universals that at some point were thought to be absolute turned out to be mere tendencies. With ever-larger language samples, typologists have found 'exceptions' to almost all cross-linguistic generalizations. It seems that grammar has very few properties that are truly universal. Therefore, functionally and cognitively oriented linguists have concentrated on the analysis of crosslinguistic tendencies, or 'statistical universals'.

The statistical approach to universals is usually motivated by theories of language change, which in turn are often motivated by usage. In accordance with this approach, B argues in the current book that language universals are the result of universal pathways of change that are grounded in usage. Therefore, the usage-based analysis of language universals often concentrates 'on the processes that create and maintain language structure, not the structures themselves' (201), or as B put it in an earlier publication, crosslinguistic research has to focus on diachronic processes because 'the true universals of language are universals of change' (2003:151).

What B adds to this hypothesis in the current discussion is that grammatical development is driven not only by general cognitive processes of language use, but also by social and cultural factors that can influence grammatical change in specific ways. As an example, she refers to a crosslinguistic study on deixis by Perkins (1992), who found that the physical and social environments of a speech community can influence the development of morphological structure in deictic expressions. Therefore, linguistic structure is shaped not only by general cognitive processes, which are in the focus of the current monograph, but also by specific cultural factors. To the extent that these factors are shared across the language users and speech communities, they generate similar linguistic structures across languages (for a similar view see Evans & Levinson 2009).

3. GENERAL DISCUSSION. The usage-based approach marks a paradigm shift, a radical departure from the structuralist and generativist traditions in linguistics. B is one of the leading usage-based linguists and her work has had a strong impact on the development of this approach. The current book provides an overview of her research, supplemented by some new data and analyses. The discussion is detailed and thorough, but not always easy to understand. Some chapters provide little background information for readers not familiar with the usage-based approach, and the structural organization of the argumentation is not always immediately obvious.

Moreover, in contrast to what is said on the book cover, the book does not present a comprehensive theory of usage-based linguistics. The bulk of the monograph is con-

cerned with B's own research and considers other usage-based research only if it is immediately relevant for B's argumentation. This does not diminish the importance of the book. On the contrary, due to the focus on the author's own research, this is a unique piece of scholarly work providing an excellent overview of B's research on grammar, usage, and grammaticalization; but it is not a systematic presentation of the usage-based approach.

If we define the notion of usage-based as in B's book, there are various other strands of research that are relevant for this approach. A comprehensive treatment of usage-based linguistics would have to integrate research from functional and cognitive linguistics, sentence processing, first language acquisition, and other research in cognitive science. In the remainder of this review article, I provide a short overview of the research areas that are relevant for a general usage-based theory of grammar and consider the relationship of this research to B's book.

B's research pertains to the functionalist and cognitivist traditions in linguistics. There is a great deal of research in this tradition that is immediately relevant for the usage-based approach. B's book is primarily concerned with the role of frequency and meaning for usage and development (see also Haspelmath 2008a,b), whereas other researchers have emphasized the importance of discourse and communication for the analysis of grammar (e.g. Givón 1995). For instance, there are studies arguing that the choice of referential expressions is determined by aspects of the ongoing discourse, which in turn can influence their development in diachronic change (e.g. Chafe 1994); and there are studies arguing that grammaticalization is affected by general communicative principles such as joint attention, which underlies the development of grammatical markers from spatial deictics (e.g. Diessel 2006). A comprehensive theory of usage-based linguistics has to combine the various strands of research in functional and cognitive linguistics into a unified theory.

Moreover, it has to pay more attention to the linear processing of language, which B considers only in connection with chunking; but there are other aspects of on-line processing that arguably influence the structure and development of grammar. Hawkins (2004) proposed a processing theory in which linguistic structure is shaped by efficiency principles in on-line sentence processing. Combining research from linguistic typology with research in psycholinguistics, Hawkins argues that many crosslinguistic generalizations about constituent order, extraction, and linguistic complexity reflect the influence of general processing principles such as 'minimize domains', which is based on the assumption that the human processor prefers linguistic structures with a short 'recognition domain'. Minimize domains provides a straightforward account for the crosslinguistic tendency to avoid center-embedding and long-distance dependencies.

Hawkins's processing theory is more closely related to the structuralist tradition in linguistics than other usage-based research, and his processing principles are not explicitly related to domain-general cognitive processes (in fact, B seems to (mis)interpret them as innate, language-specific parsing constraints; see p. 196); but there can be no doubt that Hawkins's general approach is consistent with the principle idea of usage-based linguistics, that is, that grammar is shaped by language use (or performance). What Hawkins has emphasized in his research is the importance of linearization, or on-line processing, for the organization of grammar (see also Auer 2000).

This is consistent with psycholinguistic research on sentence processing and priming, in which language use is commonly interpreted as a continuous decision-making process in which speaker and hearer have to select particular linguistic means, that is, particular lexemes and constructions, to produce or comprehend an utterance (e.g.

Bates & MacWhinney 1989). The sequential decision-making process is at the heart of language use; it determines the language users' linguistic behavior and the development of linguistic structure over time. In earlier research on sentence comprehension it was commonly assumed that on-line sentence processing is determined by structural heuristics such as 'minimal attachment' or 'late closure' (cf. Frazier 1985); but more recent research suggests that processing is driven by the language user's experience with concrete words and constructions. In the 'constraint-based approach to sentence processing' (e.g. MacDonald et al. 1994), language comprehension is based on the hearer's experience with concrete lexical expressions in particular syntactic contexts (e.g. MacDonald & Thornton 2009). In a similar vein, research on structural priming has emphasized the importance of concrete lexical expressions for language production (e.g. Pickering & Branigan 1999). Both comprehension and production involve a continuous, probabilistic decision-making process that is determined by the language users' prior experience with particular words and constructions.

Another research area that has emphasized the importance of usage and experience for the emergence of grammar is L1 acquisition. Although B and other usage-based linguists have rejected the hypothesis that language change is caused by language acquisition, usage-based research on L1 acquisition has crucially influenced the development of the usage-based approach (cf. Tomasello 2003). Like language change, language acquisition provides a window on the cognitive mechanisms that underlie the emergence of grammar. In fact, one can think of diachrony and ontogeny as two different dimensions of time in which grammar is always changing, always in flux, though under different conditions.

An early usage-based theory of language acquisition (and language processing) is Bates and MacWhinney's (1989) COMPETITION MODEL, which can be seen as a precursor of the usage-based theory of first language acquisition developed by Tomasello, Lieven, and colleagues. Like usage-based research on grammaticalization, usage-based research on first language acquisition has emphasized the importance of 'dynamic processes' for grammatical analysis. Both language change and language acquisition are driven by general cognitive processes of language use. This explains why there are so many parallels between language acquisition and diachronic change. B emphasizes that the two developments are not identical (see above), but recent reviews of the relevant literature have shown that there are striking similarities between them (cf. Diessel 2011, 2012), suggesting that children and adults rely on the same domain-general cognitive processes (although some of these processes may unfold their full potential only during the early years of life, that is, in parallel with language acquisition; cf. Tomasello 2003). If we accept the usage-based hypothesis that children and adults rely on the same cognitive processes in language use, a comparative analysis of language change and acquisition can help to determine the cognitive mechanisms that underlie the emergence of linguistic structure in the usage-based approach.

Finally, there is general, nonlinguistic research on human cognition that is potentially relevant for the usage-based approach. B is one of the few linguists who commonly refer to general research in cognitive science. Like her earlier work, the current book emphasizes the importance of exemplar-based categorization for the analysis of linguistic structure and includes a number of references to general research on analogy (59, 89) and chunking (34). But there is much more that usage-based linguistics can learn from general research in cognitive science. If the central goal of linguistic analysis is to derive language from nonlanguage, usage-based linguists have to take the nonlanguage part of the analysis more seriously. There is general research on memory and attention (e.g. Cowan

2005), on analogy and priming (e.g. Leech et al. 2007), on automatization and chunking (e.g. Schneider & Chein 2003), and on various other cognitive phenomena that are relevant for language. Moreover, cognitive scientists have developed computational models that make it possible to simulate the dynamic relationship between grammar and language usage. There is connectionist research modeling the acquisition of grammatical categories based on concrete tokens (see Elman et al. 1996), and there is research in computational linguistics in which the syntactic analysis of a novel utterance is derived from the prior processing of linguistic structure (see Bod 2003).

While some of this research is based on particular assumptions that may not be consistent with the ten theses of this review, all of the research areas mentioned in this section pursue the same general goal, namely to derive language from nonlanguage, or more specifically, to derive linguistic structure from general cognitive processes of language use. Together the various research strands constitute a new theoretical paradigm within the language sciences providing a radical alternative to the generativist and nativist theory of grammar and cognition. B's book makes an important contribution to this new paradigm, but it is primarily concerned with the author's own research. A truly comprehensive presentation of the usage-based approach would have a much broader scope.

REFERENCES

- ABBOT-SMITH, KERSTIN, and MICHAEL TOMASELLO. 2006. Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review* 23.275–90.
- AUER, PETER. 2000. On-line Syntax, oder: Was es bedeuten könnte, die Zeitlichkeit der mündlichen Sprache ernst zu nehmen. *Sprache und Literatur in Wissenschaft und Unterricht* 31.43–56.
- BATES, ELIZABETH, and BRIAN MACWHINNEY. 1989. Functionalism and the competition model. *The crosslinguistic study of sentence processing*, ed. by Brian MacWhinney and Elizabeth Bates, 3–73. Cambridge: Cambridge University Press.
- BECKNER, CLAY, and JOAN BYBEE. 2009. A usage-based account of constituency and reanalysis. *Language Learning* 59.27–46.
- BEHAGHEL, OTTO. 1932. *Deutsche Syntax: Eine geschichtliche Darstellung. Vol. 4: Wortstellung, Periodenbau*. Heidelberg: Winter.
- BELL, ALAN; JASON BRENIER; MICHELLE GREGORY; CYNTHIA GIRAND; and DANIEL JURAFSKY. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60.92–111.
- BOAS, HANS. 2003. *A constructional approach to resultatives*. Stanford, CA: CSLI Publications.
- BOD, RENS. 2003. *Beyond grammar: An experience-based theory of language*. Stanford, CA: CSLI Publications.
- BYBEE, JOAN. 1985. *Morphology*. Amsterdam: John Benjamins.
- BYBEE, JOAN. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10.425–55.
- BYBEE, JOAN. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- BYBEE, JOAN. 2002. Sequentiality as the basis of constituent structure. *The evolution of language out of pre-language*, ed. by Talmy Givón and Bertram F. Malle, 109–32. Amsterdam: John Benjamins.
- BYBEE, JOAN. 2003. Cognitive processes in grammaticalization. *The new psychology of language, vol. 2: Cognitive and functional approaches to language structure*, ed. by Michael Tomasello, 145–68. Mahwah, NJ: Lawrence Erlbaum.
- BYBEE, JOAN. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.
- BYBEE, JOAN, and PAUL HOPPER (eds.) 2001. *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.

- BYBEE, JOAN, and CAROL LYNN MODER. 1983. Morphological classes as natural categories. *Language* 59.251–70.
- BYBEE, JOAN, and JOANNE SCHEIBMAN. 1999. The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37.575–96.
- CHAFE, WALLACE. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- CHOMSKY, NOAM. 1957. *Syntactic structures*. The Hague: Mouton.
- CLARK, HERBERT H. 1996. *Using language*. Cambridge: Cambridge University Press.
- COWAN, NELSON. 2005. *Working memory capacity*. Hove: Psychology Press.
- CROFT, WILLIAM. 2000. *Explaining language change: An evolutionary approach*. Harlow: Longman.
- DABROWSKA, EWA. 2004. *Language, mind, and brain: Some psychological and neurological constraints on theories of grammar*. Edinburgh: Edinburgh University Press.
- DIESSEL, HOLGER. 2004. *The acquisition of complex sentences*. Cambridge: Cambridge University Press.
- DIESSEL, HOLGER. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17.463–89.
- DIESSEL, HOLGER. 2007. Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology* 25.108–27.
- DIESSEL, HOLGER. 2011. Grammaticalization and language acquisition. *The Oxford handbook of grammaticalization*, ed. by Heiko Narrog and Bernd Heine. Oxford: Oxford University Press, to appear.
- DIESSEL, HOLGER. 2012. Diachronic change and language acquisition. *Historical linguistics of English: An international handbook*, ed. by Alexander Bergs and Laurel Brinton. Berlin: De Gruyter Mouton, to appear.
- DRYER, MATTHEW. 1997. Why statistical universals are better than absolute universals. *Chicago Linguistic Society* 33.123–45.
- ELMAN, JEFFREY L.; ELIZABETH A. BATES; MARK H. JOHNSON; ANNETTE KARMILOFF-SMITH; DOMENICO PARISI; and KIM PLUNKETT. 1996. *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: Bradford Books/MIT Press.
- EVANS, NICHOLAS, and STEPHEN LEVINSON. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32.429–48.
- FILLMORE, CHARLES, and PAUL KAY. 1993. *Construction grammar*. Berkeley: University of California Press.
- FRAZIER, LYN. 1985. Syntactic complexity. *Natural language parsing: Psychological, computational, and theoretical perspectives*, ed. by David R. Dowty, Lauri Karttunen, and Arnold Zwicky, 129–89. Cambridge: Cambridge University Press.
- GIVÓN, TALMY. 1979. *On understanding grammar*. New York: Academic Press.
- GIVÓN, TALMY. 1995. *Functionalism and grammar*. Amsterdam: John Benjamins.
- GOLDBERG, ADELE. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- HAIMAN, JOHN. 1985. *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.
- HASPELMATH, MARTIN. 2008a. Creating economical morphosyntactic patterns in language change. *Language universals and language change*, ed. by Jeff Good, 185–214. Oxford: Oxford University Press.
- HASPELMATH, MARTIN. 2008b. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19.1–33.
- HAWKINS, JOHN A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- HOPPER, PAUL. 1987. Emergent grammar. *Berkeley Linguistics Society* 13.139–57.
- HOPPER, PAUL, and SANDRA THOMPSON. 1984. Transitivity in grammar and discourse. *Language* 56.251–99.
- LAKOFF, GEORGE. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- LANGACKER, RONALD. 1987. *Foundations of cognitive grammar, vol. 1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

- LANGACKER, RONALD. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- LEECH, ROBERT; DENIS MARESCHAL; and RICHARD P. COOPER. 2007. Analogy as relational priming: A developmental and computational perspective on the origin of a complex cognitive skill. *Behavioral and Brain Sciences* 31.357–78.
- LIEVEN, ELENA V. M.; JULIAN M. PINE; and GILLIAN BALDWIN. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24.187–219.
- LINDBLOM, BJÖRN; PETER MACNEILAGE; and MICHAEL STUDDERT-KENNEDY. 1984. Self-organizing processes and the explanation of language universals. *Explanations for language universals*, ed. by George Butterworth, Bernard Comrie, and Östen Dahl, 181–203. Berlin: Walter de Gruyter.
- MACDONALD, MARYELLEN C.; NEAL J. PEARLMUTTER; and MARK S. SEIDENBERG. 1994. Syntactic ambiguity resolution as lexical ambiguity resolution. *Perspectives on sentence processing*, ed. by Charles Clifton, Jr., Lyn Frazier, and Keith Rayner, 133–53. Hillsdale, NJ: Lawrence Erlbaum.
- MACDONALD, MARYELLEN C., and ROBERT THORNTON. 2009. When language comprehension reflects production constraints: Resolving ambiguities with the help of past experience. *Memory and Cognition* 37.1177–86.
- MURPHY, GREGORY. 2002. *The big book of concepts*. Cambridge, MA: MIT Press.
- NEWMAYER, FREDERICK J. 1998. *Language form and language function*. Cambridge, MA: MIT Press.
- PAUL, HERMANN. 1880. *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer.
- PERKINS, REVERE. 1992. *Deixis, grammar, and culture*. Amsterdam: John Benjamins.
- PICKERING, MARTIN, and HOLLY BRANIGAN. 1999. Syntactic priming in language production. *Trends in Cognitive Science* 3.136–41.
- PINKER, STEVEN, and RAY JACKENDOFF. 2005. The faculty of language: What's special about it? *Cognition* 95.201–36.
- SCHNEIDER, WALTER, and JASON M. CHEIN. 2003. Controlled and automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science* 27.525–59.
- SPERBER, DAN, and DEIRDRE WILSON. 1995. *Relevance*. 2nd edn. Oxford: Blackwell.
- TOMASELLO, MICHAEL. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- TOTTIE, GUNNEL. 1991. Lexical diffusion in syntactic change: Frequency as a determinant of linguistic conservatism in the development of negation in English. *Historical English syntax*, ed. by Dieter Kastovsky, 439–67. Berlin: Mouton de Gruyter.

Friedrich-Schiller-Universität Jena
 Institut für Anglistik/Amerikanistik
 Ernst-Abbe-Platz 8
 07743 Jena
 Germany
 [holger.diessel@uni-jena.de]

[Received 1 March 2011;
 accepted 19 July 2011]