

# Linguistics: Oxford Research Encyclopedias

## Frequency Effects in Grammar

Holger Diessel and Martin Hilpert

Subject: Morphology, Phonetics, Syntax Online Publication Date: May 2016

DOI: 10.1093/acrefore/9780199384655.013.120

### Summary and Keywords

Until recently, theoretical linguists have paid little attention to the frequency of linguistic elements in grammar and grammatical development. It is a standard assumption of (most) grammatical theories that the study of grammar (or competence) must be separated from the study of language use (or performance). However, this view of language has been called into question by various strands of research that have emphasized the importance of frequency for the analysis of linguistic structure. In this research, linguistic structure is often characterized as an emergent phenomenon shaped by general cognitive processes such as analogy, categorization, and automatization, which are crucially influenced by frequency of occurrence.

There are many different ways in which frequency affects the processing and development of linguistic structure. Historical linguists have shown that frequent strings of linguistic elements are prone to undergo phonetic reduction and coalescence, and that frequent expressions and constructions are more resistant to structure mapping and analogical leveling than infrequent ones. Cognitive linguists have argued that the organization of constituent structure and embedding is based on the language users' experience with linguistic sequences, and that the productivity of grammatical schemas or rules is determined by the combined effect of frequency and similarity. Child language researchers have demonstrated that frequency of occurrence plays an important role in the segmentation of the speech stream and the acquisition of syntactic categories, and that the statistical properties of the ambient language are much more regular than commonly assumed. And finally, psycholinguists have shown that structural ambiguities in sentence processing can often be resolved by lexical and structural frequencies, and that speakers' choices between alternative constructions in language production are related to their experience with particular linguistic forms and meanings. Taken together, this research suggests that our knowledge of grammar is grounded in experience.

Keywords: frequency, grammar, syntax, usage-based, collocations, linguistic productivity, phonetic reduction

---

# 1. Introduction to Frequency Research

Frequency is an important determinant for the acquisition and storage of knowledge. It strengthens the representation of concepts in memory and facilitates the execution of cognitive processes (Logan, 1988; Nosofsky, 1988; Schneider & Chein, 2003; Zacks & Hasher, 2002). However, although frequency is known to be an important aspect of human cognition, many linguists assume that our knowledge of grammar is largely independent of experience and practice. In fact, it is a standard assumption of the classic version of generative grammar that statistical aspects of language are irrelevant for the (innate) core of our grammatical knowledge (Chomsky, 1965).

Challenging this view, usage-based linguists and cognitive scientists have argued that all aspects of grammatical knowledge are derived from the language users' experience with frequent strings of concrete linguistic expressions. In the usage-based approach, grammar is a dynamic system consisting of fluid categories and variable constraints that are shaped by frequency of occurrence (Bates & MacWhinney, 1989; Bybee, 2006, 2010; Bybee & Hopper, 2001; Elman et al., 1996; Goldberg, 2006; Langacker, 1987, 2008; Tomasello, 2003).

Prior to this research, Labov and other variationist linguists used a statistical approach to the analysis of sociolinguistic phenomena and sociolinguistic aspects of language change (Labov, 1966, 1972); but this research was not directly concerned with frequency effects in grammar. In the usage-based approach, however, frequency is one of the main determinants for the emergence of linguistic structure and the organization of our grammatical knowledge.

Following the lead of usage-based linguists, researchers working in other frameworks, including some researchers of generative grammar, began to augment their models of grammar by a probabilistic component (e.g., Stochastic Optimality Theory; cf. Boersma & Hayes, 2001), so that today, frequency is an important concept of grammatical research in a wide range of theoretical models. However, mainstream generative grammar maintains that the core of our grammatical knowledge resides in a particular faculty of the mind that is not affected by frequency of occurrence (Newmeyer, 2003; see also Yang, 2004, who argues that statistical grammar learning can be combined with Chomsky's view of innate categories, parameters, and constraints).

The probabilistic turn in grammar research was influenced by the rise of corpus linguistics and the development of new statistical and computational tools for the analysis of quantitative data. Methodological questions of statistical modeling play a central role

in current research on grammar. However, this article concentrates on the question of how frequency affects the organization and development of morphological and syntactic structure.

There is now a large body of research indicating that frequency (or repetition) has a significant impact on sentence processing and utterance planning, and the development of linguistic structure in acquisition and change (Diessel, 2007). This article provides an overview of the research and considers the cognitive mechanisms that may account for the occurrence of frequency effects in grammar. There are various proposals in the literature as to how frequency may influence the representation and development of linguistic structure. Drawing on general research in cognitive psychology, Bybee (2006) and others have argued that exemplar theory provides a useful framework for the analysis of frequency effects in language. In the exemplar approach, categories are based on concrete tokens of experience, with overlapping properties that are grouped together in memory. Tokens with similar properties reinforce each other, creating token clusters that facilitate the categorization of novel tokens with related properties (Nosofsky, 1988). Building on this general framework, usage-based linguists have characterized linguistic categories as emergent concepts that are derived from our experience with concrete linguistic tokens, that is, words and utterances. Exemplar theory was first applied to the analysis of phonological categories (cf. Bybee, 2001; Pierrehumbert, 2003), but is now also commonly used to explain the cognitive organization and development of grammatical structure (cf. Bod, 2009; Bybee, 2010; Goldberg, 2006). Importantly, the formation of an exemplar-based category does not entail that speakers efface the memory traces of individual tokens. Instead, the mental representation of a category is largely based on the memorization of concrete speech events. Grammatical categories are thus derived from linguistic tokens and associated with particular lexical expressions, making the cognitive representation of linguistic structure much more concrete and specific than in generative theories of grammar. It is a standard assumption of this research that knowledge of grammar includes a great deal of lexically specific information about the meaning and distribution of individual expressions in particular syntactic contexts or constructions (see Diessel, 2016, for a review).

Exemplar theory provides a cognitive mechanism for the development of grammatical categories and constructions, but it does not sufficiently explain the full range of frequency effects in grammar. In the usage-based approach, grammar is commonly analyzed as a “structured inventory” of “symbolic units” (Langacker, 1987, p. 57), which are mutually associated through various types of links that create a system of linguistic structures that one might characterize as a network (see Diessel, 2015, for a recent discussion of the network metaphor of usage-based grammar; see also Hilpert, 2014, pp. 50–73). If we think of grammar as a network of symbolic units, frequency does not only

strengthen the cognitive representations of linguistic elements in memory (as suggested by exemplar theory) but also reinforces the associative connections between them. Other things being equal, the more often linguistic elements occur together in language use, the stronger is the associative bond between them in memory. The psychological mechanism that underlies the language users' knowledge and processing of co-occurrence patterns is automatization (Logan, 1988), also known as "entrenchment" (Langacker, 1987). Like exemplar learning, automatization is a general psychological mechanism that is crucially driven by frequency of occurrence and not restricted to language (Schneider & Chein, 2003). However, in contrast to exemplar learning, automatization is not concerned with the emergence and organization of categories, but with the processing of associative connections between concepts and category features (see Diessel, 2016, for a comparative discussion of exemplar learning and automatization).

In what follows, we consider the influence of exemplar learning and automatization on the cognitive organization of grammar. The research we review comes from a wide range of different subfields in linguistics and psychology and is not restricted to research in the usage-based model. Overall, we discuss twelve linguistic processes that display frequency effects in the use and development of linguistic structure: (1) the emergence of collocations and; (2) syntactic constituents; (3) the interaction between lexemes and constructions; (4) the productivity of linguistic schemas; (5) the ability of language users to assess the grammaticality of novel linguistic forms; (6) the occurrence of phonetic reduction and coalescence in language change; (7) the segmentation of the speech stream and; (8) the extraction of syntactic categories in L1 acquisition; (9) the maintenance of frequent linguistic strings under pressure from analogy; (10) the choice between alternative structures in language production; (11) the processing of the unfolding sentence in language comprehension; and (12) the flagging or marking of infrequent forms.

## 2. The Emergence of Collocations

Collocations are sequences of two or more words that frequently co-occur and often develop into linguistic units in their own right (*all of a sudden, I wonder if*). While the occurrence of multi-word sequences can be semantically motivated, semantic criteria alone are not sufficient to explain the existence of collocations (Taylor, 2012; see 146–166 for a recent discussion). Collocations are strings of multi-word expressions that have become conventionalized by frequency or repetition (Erman & Warren, 2000; Wray, 2002). For instance, every native speaker of English is familiar with the expressions *strong tea*

and *powerful computer* and knows that, although the adjectives *strong* and *powerful* are semantically related, they are not interchangeable in these expressions; that is, while the strings *powerful tea* and *strong computer* are consistent with general semantic principles, native speakers know that these strings are unusual, or nonidiomatic (Manning & Schütze, 1999, pp. 153–157). Or consider the expression *unmitigated disaster* (Taylor, 2012, chapter 7). Although *disaster* is much less frequent than semantically related nouns such as *accident* and *mishap*, people associate *unmitigated* with *disaster*, rather than with *accident* or *mishap*, because they have encountered the string *unmitigated disaster* much more frequently than the strings *unmitigated accident* or *unmitigated mishap* (see Taylor, 2012, pp. 158–161 for data and discussion). Most research on collocations is based on corpus data, but there is also experimental evidence for the hypothesis that frequent word strings are stored and processed as conventionalized units (Arnon & Snider, 2010).

In the generative approach, collocations are treated as a marginal phenomenon that is excluded from grammatical theory; but in other theoretical frameworks, it is widely assumed that language users' knowledge of these expressions cannot be ignored in grammatical analysis.

Collocations restrict speakers' linguistic choices and often develop idiosyncratic properties that are not immediately predictable from the properties of their components. There is a continuum of idiosyncrasy, ranging from frequent multiword expressions that are fully compositional and licensed by general grammatical patterns or rules (e.g., *I am happy*) to highly idiomatic expressions that exhibit idiosyncratic semantic properties and deviate from other grammatical forms (e.g., *all of a sudden*). The existence of this continuum challenges the traditional distinction between "grammatically derived" and "idiomatic," or "grammar" and "lexicon," and has played a central role in the development of Construction Grammar (Fillmore, Kay, & O'Connor, 1988). If there is no clear-cut boundary between derived and stored expressions, it seems reasonable to assume that our grammatical knowledge includes a large number of prefabricated strings and that the acquisition of grammar crucially involves "sequence learning" (Ellis, 1996, p. 92).

What is more, the omnipresence of prefabricated strings affects our knowledge of grammatical categories. Argument structure, for instance, is crucially influenced by the existence of collocations and idioms. As Thompson and Hopper (2001) point out, speakers of English know a whole array of "verb-object compounds" consisting of light (transitive) verbs and nonreferring nouns, as in examples (1, a-d).

(1)

- a. I'll **have fun**.
- b. Your clues **make no sense**.
- c. I need to **get sleep** over the weekend.
- d. **Wait a minute**.

At the surface, the examples in (1) have the structure of an ordinary transitive clause; but since the object NP has lost its referential meaning and merged with the preceding verb to a new semantic unit, these structures are semantically equivalent to intransitive clauses, indicating that basic syntactic categories such as argument structure and transitivity are immediately affected by our knowledge of collocations and idioms. According to Thompson and Hopper, the traditional theory of argument structure ought to be replaced by a “probabilistic theory” of verb-noun combinations that reflect the language users’ experience with particular strings of multiword expressions.

### 3. The Emergence of Syntactic Constituents

Closely related to this analysis of argument structure is the usage-based view of constituency. In structuralist and generative approaches, constituent structure is derived from a small set of syntactic categories (e.g., DET, N, NP, PP) that are combined by general phrase structure rules (e.g., NP → DET N) into discrete, hierarchical configurations, often represented as syntactic tree structures. Challenging this view, Bybee (2002, 2010) has argued that phrase structure is emergent from language users’ experience with frequent strings of linguistic elements and hence probabilistic, rather than discrete (see also Bybee & Scheibman, 1999).

In natural language processing (i.e., computational approaches to syntax), phrase structure analysis is often supported by a stochastic component in which individual rules are assigned a probability value that reflects the relative frequency of a particular phrase structure rule in a corpus (e.g., Bod, 2009; Jurafsky, 1996). Consider, for instance, the following (made-up) examples:

Table 1. Stochastically Enriched Phrase Structure Rules

Number	Phrase structure rule	Example	Probability value
1	VP → V	<i>(He) died.</i>	0.39
2	VP → V NP	<i>(He) saw this movie.</i>	0.41
3	VP → V NP NP	<i>(He) gave John the key.</i>	0.08
4	VP → V NP PP	<i>(He) send a letter to his new employer.</i>	0.12
Total			1.00

It is well-known that syntactic sequences are often ambiguous and therefore difficult to parse (without semantic information); but if phrase structure rules are enriched by a probabilistic component, as the VP-rules in Table 1, the parser can easily compute the most probable analysis of an ambiguous string such as *v\_NP\_PP*, in which the prepositional phrase can function either as an oblique argument (e.g., *He put the book on the table*) or as an attribute of the preceding noun phrase (e.g., *He saw the book on the table, but did not notice the one on the shelf*) (Jurafsky, 1996).

The statistical approach to syntactic parsing has greatly improved the results of automated language processing (Jurafsky & Martin, 2000; Manning & Schütze, 1999); but it is based on a rather traditional view of syntax consisting of discrete categories and rules that are defined prior to grammatical analysis. Challenging this approach, usage-based linguists have argued that phrase structure is an emergent phenomenon shaped by two general cognitive processes: (a) semantic coherence and (b) automatization or “chunking” (Bybee, 2010).

It is a well-known tendency across different languages that speakers tend to place semantically related elements next to each other (Behaghel’s first law, cf. Diessel, 2015, p. 315). Langacker (2008, p. 207) characterizes “classic constituents” as “a particular kind of conceptual grouping” consisting of a “trajector” and a “landmark.” However, in addition to conceptual factors (i.e., semantic coherence), syntactic constituents are influenced by automatization. Like strings of lexical expressions, strings of syntactic categories become associated with each other if they are frequently processed. One can think of syntactic constituents as schematic collocations in which emergent grammatical categories are bound together to a processing unit. The more often two or more categories occur

together, the stronger is the bond between them. On this account, constituency forms a continuum ranging from structures that are closely related (e.g., determiner and noun) to structures that are only loosely associated with each other (e.g., verb and manner adverb) (see Bybee, 2002; Bybee & Scheibman, 1999; Bybee, 2010 for discussion).

## 4. The Interaction Between Lexemes and Constructions

Like other aspects of grammar, phrase structure exhibits frequency effects that are ultimately determined by speakers' experience with particular lexical expressions. Most usage-based linguists conceive of syntactic constituents as constructions—conventionalized sequences of linguistic elements combining a particular form with a particular meaning or function. Constructions vary on a scale of schematicity ranging from collocations and idioms to highly schematic representations including slots for particular lexical expressions (Langacker, 2008, p. 19). Consider, for instance, the constructional schema of the English passive construction, consisting of an initial noun phrase, denoting a patient or theme, a periphrastic verb form, including an inflected form of the verb *be* and a past participial, and optionally a *by*-phrase, denoting the agent or experiencer of the activity expressed by the verb (cf. 2, a–d).

(2)

- a. The factory was built by the company.
- b. The soup was cooked before dinner.
- c. The house was destroyed by the wind.
- d. The book was written by a great poet.

The passive construction can occur with a wide range of (transitive) verbs, for example, *build*, *cook*, *destroy*, and *write*. However, the co-occurrence of transitive verbs and passive voice is not random. As Gries and Stefanowitsch (2004) have shown, individual verbs such as *use*, *involve*, and *publish* are more frequent in the passive construction than one would expect based on their overall frequency in a corpus. Following Goldberg (1995), it is commonly assumed that the co-occurrence patterns of lexemes and constructions are semantically motivated; that is, lexemes and constructions “fuse” if they are semantically compatible with each other (Goldberg, 1995, p. 50). However, a number of studies have pointed out that the co-occurrence patterns of lexemes and constructions are not fully predictable from general semantic criteria (e.g., Boas, 2003). There is, for instance, no

obvious semantic reason why the verbs *use*, *involve*, and *publish* are more frequent in the passive construction than statistically expected, and why other verbs, such as *think*, *say* and *want*, are predominantly found in the corresponding active construction (Gries & Stefanowitsch, 2004). However, notwithstanding semantic criteria, native speakers know these co-occurrence patterns because of their experience with particular words and constructions (see Diessel, 2015, for discussion). In the generative approach, lexical expressions are freely inserted under the terminal nodes of syntactic phrase structure trees; but in the usage-based approach, constructions are associated with individual lexical expressions that are more frequent in particular structural positions (or “slots”) than one would expect on statistical grounds.

## 5. The Productivity of Linguistic Schemas

The associative connections between lexemes and constructions play an important role in the analysis of linguistic productivity. In the generative approach, syntactic productivity is commonly defined by the (recursive) application of general phrase structure rules; but in the usage-based approach, linguistic productivity is usually defined as the likelihood that a constructional schema will be applied to new lexemes (Bybee, 2010, p. 94; Langacker, 2000, p. 26). Each slot of a construction is associated with a class of lexical expressions that have appeared in these positions on earlier occasions; but the co-occurrence of lexemes and constructions is not restricted to established patterns. Speakers can extend the use of a constructional schema to novel expressions, for instance, by borrowing a lexeme from another language or using a given word in a novel (syntactic) context (e.g., *She smiled herself an upgrade*; Goldberg, 2006, p. 6). The extension of constructional schemas to novel expressions is based on structure mapping or analogy, which is crucially influenced by similarity (Gentner, 1989). There is evidence from a number of studies that constructional schemas are applied to novel items if these items are semantically and/or formally related to expressions that are already licensed by a particular schema. Consider, for instance, the following examples from Goldberg (1995) and Boas (2003).

(3)

- a. She sneezed the napkin off the table.
- b. The wind blew the leaves off the tree.
- c. ??Frank wheezed the napkin off the table.

The sentence in (3a) is a frequently cited example of “coercion,” in which the intransitive verb *sneeze* is interpreted as a causative verb in the context of the resultative construction (Goldberg, 1995, p. 5). Coercion is an item-specific process that involves the unusual combination of a particular lexeme and construction licensed by analogy. According to Boas (2003), *sneeze* is more easily accommodated to the resultative construction than *wheeze*, because *sneeze* is semantically more similar to verbs such as *blow* that are commonly used in the resultative construction. In accordance with this view, historical linguists have shown that constructional change typically proceeds in an item-based fashion that is crucially driven by the similarity between individual lexical expressions (see Hilpert, 2013; see also De Smet, 2012; Israel, 1996; Traugott & Trousdale, 2013).

Frequency affects the productivity of constructional schemas in two important ways. First, lexical expressions that are frequently used in a specific grammatical pattern (and are therefore strongly associated with it) come to be represented as lexical prototypes for that grammatical pattern. As a result, they are more likely to license the extension of that grammatical pattern to a semantically related lexeme than infrequent expressions that are less strongly associated with it. In concrete terms, it is very likely that new extensions of the English ditransitive construction will be modeled on the verb *give*, but not on the verb *deny*. And second, the productivity of a slot varies with type frequency and the presence of hapax legomena (i.e., types that are registered only once in a given body of data). Assuming an average amount of (dis)similarity between expressions of a particular word class, the more lexical types are associated with a particular position in a constructional schema, the less specific are the semantic and formal constraints of this position, so that constructional schemas with high type frequency are often highly abstract, which facilitates their extension to novel expressions (cf. Bybee, 1985, 1995; Goldberg, 1995, pp. 134–137). A high ratio of types that occur only once indicates that speakers create new extensions on a regular basis (cf. Baayen, 1993).

## 6. The Grammaticality of Linguistic Forms

Productivity is an important aspect of language use that refers to the speaker’s ability to produce novel utterances; but this ability is not unlimited. Native speakers know that certain strings of linguistic elements are unnatural or outright ungrammatical. In the generative approach, grammaticality is a discrete concept that researchers use to identify grammatical rules; but in many other frameworks, grammaticality is a gradient notion that is grounded in the language users’ experience with particular lexemes and

constructions. In this approach, novel sentences can be more or less grammatical, depending on their relationship to the language users' linguistic knowledge or past linguistic experience. Since linguistic experience varies across speakers, it does not come as a surprise that grammaticality judgments correlate with social parameters, such as the educational background or profession of individual speakers (see Dąbrowska, 2012, for a review of relevant research). However, the cognitive mechanisms that underlie grammaticality judgments are the same across speakers. They are determined by the same factors as productivity, namely similarity and frequency.

To simplify, novel sentences appear to be grammatical if they correspond to established co-occurrence patterns with local or minor analogical extensions. Of course, similarity can concern different aspects of language, such as linear order, morphology, or semantics, making it difficult to predict the degree of grammaticality from a general notion of linguistic similarity. However, notwithstanding the difficulty to define similarity, there is evidence that grammaticality judgments are crucially influenced by the amount of overlap or similarity between a novel sentence and stored grammatical patterns.

The role of frequency has also been emphasized in research on L1 acquisition, which seeks to explain why children do not acquire “an overly general grammar” (Bowerman, 1988). Preschool children overgeneralize grammatical schemas or rules, producing strings of linguistic elements that are not acceptable or ungrammatical in adult grammar (e.g., *Don't giggle me*); but these structures disappear in the course of language development. How do children learn to constrain the use of grammatical patterns and to avoid overgeneralization errors? A number of studies have argued that narrowly defined semantic verb classes play an important role in the constraining of grammatical constructions (Pinker, 1989). However, in addition to semantic factors (i.e., semantic similarity), it is the frequency of particular co-occurrence patterns that shapes the child's growing ability to avoid the overuse of grammatical patterns. Other things being equal, children are more likely to overextend the use of a constructional schema to an infrequent word, rather than to a frequent one. For instance, Brooks, Tomasello, Dodson, and Lewis (1999) showed that preschool children are relatively more open towards extending the transitive construction (e.g., *He cut the rope*) to an infrequent intransitive verb such as *vanish* (e.g., *He vanished the rabbit*), as opposed to extending it to a frequent intransitive verb with the same meaning such as *disappear* (e.g., *He disappeared the rabbit*). This indicates that grammaticality is not just a matter of similarity but also of frequency or entrenchment (see also Ambridge, Pine, Rowland, & Young, 2008; Stefanowitsch, 2008).

## 7. Phonetic Reduction and Coalescence

One of the best known and most intensively analyzed effects of frequency is phonetic reduction (e.g., Bell et al., 2003; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Bybee, 1985, 2001; Gahl, Yao, & Johnson, 2012; Jurafsky, Bell, Gregory, & Raymond, 2001). The reduction effect of frequency can be observed in both synchronic language use and diachronic language development. However, since phonetic reduction also correlates with other parameters of language use, such as the linguistic context, the speech rate, and the speaker's age, it is not easy to determine the precise effect of frequency on phonetic reduction. Using multi-factorial regression models, Bell et al. (2009) showed that there is a strong negative correlation between frequency of occurrence and the degree of phonetic reduction if all other factors are controlled for. However, the correlation is not uniform across expressions. For instance, while (simple) word frequency correlates with the degree of phonetic reduction in content words, it does not seem to correlate with the frequency of function words. Specifically, Bell et al. observed that regardless of the linguistic context, frequent content words are more strongly reduced than infrequent ones, whereas function words are only phonetically reduced if their occurrence is predictable from the linguistic context (and regardless of their total frequency) (cf. Bell et al., 2003; Jurafsky et al., 2001).

Note that there is no consensus among researchers as to why speakers tend to reduce frequent (strings of) linguistic elements. According to Bybee (2001), phonetic reduction is primarily caused by the automatization of articulatory gestures; but other researchers have claimed that it is the greater predictability of frequent expressions that leads speakers to reduce the amount of articulatory effort (e.g., Jurafsky et al., 2001). The two factors, the automatization of speech gestures and the predictability of co-occurring words, are not mutually exclusively and may complement each other (Bybee, 2010, p. 38–43); but more research is needed to understand the cognitive and neuromotor processes that lead to phonetic reduction in speech production.

Phonetic reduction in language use can have long term effects on language development that are immediately relevant for the organization of grammar. It is well known that grammatical markers are commonly derived from frequent content words (or spatial deictics; Diessel, 2012A), and that this development typically involves phonetic reduction. In contrast to nouns and verbs (and spatial deictics), function words are commonly unstressed and formally reduced to the point that speakers are usually unable to identify grammatical markers if they are sliced out of context and presented in isolation (Pollack & Pickett, 1964).

Although research on grammaticalization has focused on individual grammatical items, it must be emphasized that grammaticalization generally concerns strings of linguistic expressions rather than isolated words (e.g., *be going to*, *in front of*). Grammaticalization is a complex phenomenon involving both formal and semantic changes. While these changes are driven by several cognitive processes, such as metaphor, analogy, and pragmatic inference, most linguists agree that frequency (or automatization) is the main determinant of phonetic reduction in grammaticalization (cf. Bybee, 2003; Krug, 2003).

In the course of this development, grammatical markers often lose their status as independent words and merge with neighboring expressions. Coalescence is a frequent phenomenon of language change that accounts for the existence of bound morphemes. There is a well-known diachronic cline, leading from independent (function) words to (grammatical) affixes, via clitics, that correlates with “string frequency” (Krug, 1998), that is, the frequency of neighboring expressions. In this view, morphology is an emergent phenomenon derived from frequent linguistic sequences, which Givón (1971) epitomized in the slogan *Today’s morphology is yesterday’s syntax*.

## 8. Segmentation of Phonetic Sequences

All of the processes we have considered thus far involve the development of automated processing units. However, interestingly, in language acquisition, frequency also plays an important role in the segmentation of automated sequences. When children are born, they have no concept of morpheme, word, or phrase, and thus have to “unpack” the phonetic sequences they encounter in the ambient language. This is one of the most fundamental tasks of (early) language acquisition and a prerequisite for grammar learning (Jusczyk, 1997).

There are two important types of cues children use to break into linguistic structure. First, there are phonetic cues: pauses, intonation, and phonotactic constraints that help the child to divide phonetic sequences into particular units; and second, there are distributional cues, or distributional regularities, that are potentially available to identify the boundaries between particular words and phrases (cf. Jusczyk, 1997).

In a seminal study, Saffran, Aslin, and Newport (1996) have demonstrated that young children are very sensitive to distributional regularities in phonetic sequences. Using four meaningless nonce words, they constructed a set of uninterrupted and prosodically unmarked strings of words and exposed 8-month-old infants to these sequences. The four nonce words were composed of three CV-syllables, *tupiro*, *golabu*, *bidaku*, and *padoti*,

that were spliced together in random order and without intonational pauses. After the infants had listened to these sequences for two minutes, they were tested under two conditions. In the first condition, they had to listen to a new string of the same four nonce words in random order; but in the second condition, the four nonce words were decomposed into syllables and children had to listen to a random string of syllables (rather than a string of words). Since the infants were familiar with the four nonce words, Saffran et al. hypothesized that, if children are sensitive to distributional regularities in phonetic sequences, they would recognize the difference between the two conditions. In accordance with this hypothesis, the researchers found that the infants of their study listened longer to the string of syllables than to the string of words. Since the experimental stimuli did not include any semantic or prosodic features, it must have been the different distributional properties of words and syllables that led to these responses. Specifically, Saffran et al. hypothesized that the infants of their study recognized that the transitional probabilities between syllables in the word-condition are much higher than those in the syllable-condition, suggesting that statistical regularities in the ambient language might play a central role in the segmentation of the speech stream (see Aslin & Newport, 2012, for a comprehensive review of subsequent research on this topic; see also Siegelman & Frost, 2015, who argue that there are great individual differences between children's sensitivity to conditional probabilities).

## 9. Extraction of Syntactic Categories

Inspired by this finding, researchers began to explore the role of distributional learning in the acquisition of grammar. According to Chomsky, the language children experience is too simple, too fragmented, and too inconsistent in order to learn grammatical categories from experience alone. The so-called "argument from the poverty of the stimulus" has played a key role in the theory of linguistic nativism (Pinker, 1989). However, a number of corpus studies have shown that child-directed speech is much more regular and systematic than commonly assumed in generative theories of language acquisition. In one of these studies, Redington, Chater, and Finch (1998) examined the bigram statistics of the one thousand most frequent words in the ambient language of the entire English component of the CHILDES database. Using a series of computational experiments, they showed that a hierarchical cluster analysis of bigram statistics groups the words of the ambient language into a structured set of word classes that corresponds very closely to the traditional inventory of word class categories. This indicates that children could, in principle, extract grammatical categories such as noun, verb, and preposition from a distributional analysis of the ambient language (and without the support of an innate language faculty). Related research by Mintz, Newport, and Bever

(2002) and Monaghan, Chater, and Christiansen (2005) improved the results of the Redington study by augmenting the analysis with information about phrasal boundaries and phonological features.

Complementary to this line of research, other scholars examined children's ability of statistical grammar learning by means of experimental methods. For instance, Marcus, Vijayan, Rao, and Vishton (1999) conducted an experiment in which they taught 7-month-old infants two different "sentence types," defined here as short patterns of linguistic structure. One group of children learned sentences that followed an ABA pattern, and another group of children learned sentences that followed an ABB pattern. The sentences were instantiated by monosyllabic nonce words such as *ga*, *to*, and *ni*, yielding strings such as *ga-to-ga* (i.e., ABA) or *li-ti-ti* (i.e., ABB). After training, the researchers replaced the words of the training phase with novel expressions and exposed the infants to a new battery of sentences; but now, all children were exposed to both sentence types: the one they had heard during training and the one they had not heard before. Although the words of the test sentences were entirely new to the children, they recognized the different distributional patterns, indicating that they had generalized across the words of these sentences. More specifically, the children had extracted schematic representations of linguistic structure from strings of phonetic tokens, which Marcus et al. interpreted as evidence for the acquisition of a syntactic rule, but which can be analyzed as the extraction or emergence of a constructional schema from linguistic tokens. Other experimental research by Gómez and Gerken (1999) and Gerken (2004) confirmed the results of this study, supporting the general conclusion that statistical learning plays an important role in the acquisition of syntax (see Aslin & Newport, 2012, for a review of this research).

## 10. Maintenance Under Pressure From Analogy

Since frequent strings of linguistic elements are deeply entrenched in memory, they are more resistant to analogy (or structure mapping) than infrequent ones (cf. Bybee, 1985, 2010). Analogy plays an important role in both language acquisition and diachronic language change. Both adults and children are prone to accommodate linguistic elements to general structural patterns by analogy (see Diessel, 2012B, for a comparison of structure mapping in L1 acquisition and language change). For instance, there is a general tendency in English to regularize irregular verb forms such as *keep*, *blow* and *hit*. Old English had about 300 irregular verbs, but many of them were regularized in the development from Old to Modern English. The same tendency of regularizing irregular verbs occurs in L1 acquisition. At the age between 3;0 and 4;0, children produce

overextension errors such *keeped*, *blowed*, and *hitted* (cf. Diessel, 2012B). Interestingly, the occurrence of children's errors seems to be determined by the same factors as the diachronic development of irregular past tense forms. Two factors are important: The first factor relates to type frequency: irregular verbs that are phonetically associated with a specific past tense schema, or a larger class of irregular verbs (e.g., *sing-sang*, *ring-rang*, *shrink-shrank*, *sink-sank*, etc.), are less likely to be leveled by analogy than irregular verbs that are not (or only loosely) associated with a phonetic verb class (e.g., *fall-fell*, which lacks "companions"). The second factor is frequency of occurrence: frequent irregular verbs are less likely to be regularized than infrequent ones because they are more strongly represented in memory and hence not so easily changed by structure mapping (cf. Bybee & Slobin, 1982).

The interaction between entrenchment and analogy is not restricted to morphology. The same factors influence the development of syntactic schemas. For instance, a number of studies have argued that the development of negated sentences in Early Modern English followed a trajectory that is crucially determined by the frequency of individual expressions (Bybee, 2010, p. 6971; Krug, 2003; Tottie, 1991). Old English had several strategies to form negative sentences, such as a negative particle, *ne*, that preceded the verb, and negative indefinite markers consisting of *ne* and a pronoun or quantifier that followed the verb. The latter provided the source of the present day negative marker *not*, consisting of *ne* plus *wiht* meaning "not at all" (Tottie, 1991). In Early Modern English, it became increasingly less common to express negation by a postverbal negative indefinite marker, and a new pattern emerged in which negation is expressed by an auxiliary or modal plus *not*, followed by the main verb. The development proceeded in an item-specific fashion that correlates with verb frequency. As Tottie (1991) and Krug (2003) have demonstrated based on diachronic corpus data, in Early Modern English postverbal *not* was especially frequent with the present day auxiliaries and modals, such as *have*, *be*, *can*, and *must*, which have preserved the old pattern of postverbal negation, whereas all other verbs are now negated by a preverbal negative marker and an auxiliary, as in *haven't VERB*, *isn't VERB*, *won't VERB*.

## 11. Utterance Planning and Production

Language is a sequential activity in which speaker and hearer are forced to make rapid online decisions. Speakers have to select particular words and constructions to produce an utterance, and hearers have to link the phonetic signals they receive to particular

concepts of their linguistic knowledge. Both utterance planning and sentence comprehension are influenced by frequency of occurrence.

Very often, there are alternative ways of expressing a particular intention or meaning. What determines the speaker's choice of linguistic means in language production? One factor that has a significant impact on speaking is priming, that is, the activation of a cognitive circuit that facilitates the subsequent activation of a related circuit. There is evidence from a wide range of studies that utterance planning and production are crucially influenced by the linguistic elements that have been activated in the previous discourse. The effects of lexical priming (e.g., *honey* priming *bee*, *doctor* priming *nurse*) have been well known for a long time, but there is now also a large body of research indicating that priming affects not only lexical access, but also the speaker's selection of morphosyntactic structures. If speakers can choose between alternative structures, and if one of these structures has been previously activated, they are likely to reuse this structure in the unfolding discourse (see Pickering & Ferreira, 2008, for a review of research on structural priming).

Like exemplar learning and automatization, priming concerns the activation status of linguistic elements in memory; but since priming is commonly characterized as a short-term phenomenon of working memory, it is not immediately relevant for the analysis of frequency effects in grammar. There is evidence, however, that language production is influenced not only by the transient activation patterns of working memory, or priming, but also by the speaker's long-term linguistic knowledge. A number of studies have shown that frequent co-occurrence patterns facilitate utterance planning and speech production. Specifically, these studies suggest that speakers' linguistic choices between alternative structures are predictable from their experience with particular words and constructions.

For instance, Bresnan et al. (2007) conducted a corpus study in which they examined the so-called dative alternation, the alternation between the double-object construction (e.g., *Peter gave John the key*) and the *to*-dative construction (e.g., *Peter gave the key to John*). Using logistic regression models, they showed that the speaker's choice between the two constructions is predictable, with a high degree of accuracy, from a set of linguistic features that tend to co-occur in one or the other of the two constructions in a corpus. For instance, given a "known" and "animate" recipient and an "unknown" and "inanimate" theme, chances are very high that speakers select the double-object construction, rather than the *to*-dative, in order to express a ditransitive scene (i.e., a scene involving transfer of an object from an actor to a recipient).

A similar regression study was conducted by Diessel (2008), who showed that the alternation between pre- and postposed temporal adverbial clauses (e.g., *After it began to*

*rain, they left* vs. *They left, after it began to rain*) is statistically predictable from three general criteria that influence the cognitive processes of utterance planning and production, and that speakers know from their past linguistic experience: the iconicity of clause order, the relative length of main and adverbial clauses, and the occurrence of a causal or conditional interpretation implied by the temporal clause (see also Diessel, 2005; Wiechmann & Kerz, 2013).

## 12. Sentence Processing and Structural Ambiguity Resolution

Like utterance planning, sentence comprehension is crucially influenced by frequency of occurrence. One of the earliest and most influential studies of sentence processing that emphasized the importance of frequency for sentence comprehension is Bever (1970). Drawing on data from a series of experiments, Bever argued that there is a strong tendency in English to interpret a preverbal NP as the agent of the sentence. Since basic declarative sentences tend to express the agent prior to all other argument roles, non-canonical sentence types can incur additional processing costs if they deviate from the expected pattern. Passive sentences, for instance, cause prolonged reading times, compared to basic declarative sentences, because they include the patient or theme prior to the agent, which is only optionally expressed in a postverbal *by*-phrase (cf. *the thief that was chased by the police*).

The same analysis applies to complex sentences with reduced relative clauses, as Bever's famous example *The horse raced past the barn fell*. Assuming that the clause-initial NP of this sentence serves as the agent of the subsequent verb, there is a strong tendency to interpret the verb *raced* as the past tense form of a simple (in)transitive clause; but since this interpretation is not consistent with the verb *fell* at the end of the sentence, the listener is forced to revise the initial parse. This explains, according to Bever, why reduced relative clauses can lead the hearer down a garden path.

Building on this analysis, more recent research has shown that the processing costs of reduced relative clauses are crucially influenced by lexical frequencies. For instance, given that reduced relative clauses evoke a passive interpretation, Trueswell (1996) hypothesized that these structures are easier to process with verbs that are frequently used in passive voice than with verbs that are primarily used in active voice. Using a self-paced reading task, he compared the reading times that occurred in response to two different stimuli: reduced relative clauses, including verbs such as *select*, that are

frequently used in passive voice (cf. 4a); and reduced relative clauses, including verbs such as *search*, that are primarily used in active voice (in the past tense) (cf. 4b).

(4)

- a. The recipe selected by the judges did not deserve to win.
- b. The room searched by the police contained the missing weapon.

As predicated, Trueswell found that reduced relative clauses that include a verb such as *select* cause significantly fewer processing difficulties than reduced relative clauses that contain a verb such as *search*, which is only rarely used in passive voice, suggesting that the language users' experience with particular verb forms (i.e., active vs. passive) has a significant impact on the interpretation of this construction (see also MacDonald, 1994).

Parallel results have been obtained in a large number of other processing studies investigating other types of ambiguities (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Seidenberg & MacDonald, 1999; Spivey-Knowlton & Sedivy, 1995). Taken together, this research suggests, in accordance with Bever's classic study, that sentence comprehension is guided by language users' experience with statistical co-occurrence patterns of lexemes and constructions (see also Wiechmann, 2008).

## 13. Typological Markedness and Morphological Flagging

Finally, frequency is commonly evoked to explain cross-linguistic patterns of markedness. In linguistic typology, the term *markedness* refers to structural asymmetries in morphological paradigms, which were first described by Greenberg (1966). Building on Greenberg, Croft (2003) distinguished several types of typological markedness; but here we concentrate on "structural markedness," which is perhaps the most common kind of typological markedness.

Typologists agree that structural asymmetries in morphological paradigms correlate with frequency, and that frequency of occurrence must have played some role in the diachronic evolution of markedness patterns; but there is no consensus among typologists as to how exactly frequency or experience has influenced the development of these asymmetries.

The classic example of structural markedness is nominal plural. Across languages, plural nouns are much more likely to be marked by an affix than singular nouns (cf. Engl. *car* vs. *car-s*). There are languages in which both singular and plural nouns occur with a particular marker (cf. Zulu *umu-ntu* “SG-person” vs. *aba-ntu* “PL-person”), and other languages in which both singular and plural nouns are unmarked (cf. Minor Mlabri *ʔεεw* “child” vs. *ʔεεw* “children”); but there seems to be no language in which singular nouns are generally combined with a number affix, whereas plural nouns are unmarked (Croft, 2003, pp. 88–89). What is attested in some languages is that individual nouns take a number affix in the singular and no marker in the plural; but this is always a local phenomenon, restricted to nouns that typically refer to entities that appear in groups or pairs, such as nouns for certain types of animals (e.g., sheep, bees) or nouns for certain body parts (e.g., eyes, ears) (Tiersma, 1982). Apart from nominal plural, morphological asymmetries are also commonly found in various other grammatical categories. For instance, across languages, the subject is less likely to occur with a case affix than the object or an adverbial; active verb forms are less likely to occur with a particular affix than verbs in passive voice; and affirmative sentences are less likely to include a particular marker than negative sentences (cf. Croft, 2003; Greenberg, 1966).

How do we account for these asymmetries? There are several explanations. Some researchers have argued that structural markedness patterns are motivated by a general principle of economy. On this account, frequent category members tend to be unmarked because it is more economical to express grammatical distinctions by marking the infrequent category member than the frequent one. Other researchers have argued that structural markedness patterns can be explained by phonetic reduction. On this account, frequent category members are unmarked because automatization leads to the erosion of phonetic material and the subsequent loss of grammatical markers. However, while both of these explanations are not implausible, there is little empirical evidence to support them. In fact, Haspelmath (2008) argued that neither economy nor phonetic reduction is sufficient to explain cross-linguistic asymmetries in the marking of grammatical categories. According to Haspelmath, frequency correlates with morphological marking primarily because frequency is an important determinant of what people expect in particular situations. Specifically, he argues that since infrequent category members are less expected than frequent ones, they need some kind of morphological flagging. On this account, structural markedness patterns have evolved from the need to signal that the present element deviates from the (expected) default in order to facilitate communication.

## 14. Significance of Frequency Effects

To summarize, this paper has given an overview of frequency effects in grammar and grammatical development. The research that we have reviewed supports a view of linguistic knowledge in which frequency of use is a fundamental determinant of grammatical knowledge. This view goes against several long-standing traditions in linguistics: The Saussurean dichotomies of the linguistic system (*langue*) vs. language use (*parole*), and language development (diachrony) vs. the current state (synchrony), which were important cornerstones of linguistic structuralism, still inform many branches of contemporary linguistics, including generative linguistics, but also less formal approaches. As we hope to have shown, there is now a substantial body of empirical work that calls these dichotomies into question. Speakers' knowledge of grammar is fundamentally grounded in their experience with concrete words and utterances, which crucially involves frequency of occurrence, so that a crisp distinction of system and use cannot be upheld. The acquisition and diachronic development of linguistic structure is shaped by general cognitive processes such as exemplar learning, automatization, and analogy, so that synchrony and diachrony cannot be fully understood in mutual isolation. All of the processes we have discussed are crucially influenced by frequency. In summary, then, frequency is not just a performance phenomenon, distinct from mental grammar. Rather, the frequency with which linguistic forms are experienced is at the heart of our grammatical knowledge.

### Further Reading

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review* 23, 275–290.

Bod, R., Hay, J., & Jannedy, S. (2003) (Eds.). *Probabilistic linguistics*. Cambridge, MA: MIT Press.

Bybee, J. (2007). *Frequency and the organization of language*. Oxford: Oxford University Press.

Diessel, H. (2009). On the role of frequency and similarity in the acquisition of subject and non-subject relative clauses. In T. Givón & M. Shibatani (Eds.), *Syntactic complexity* (pp. 251–276). Amsterdam: John Benjamins.

Divjak, D., & Gries, S. T. (Eds.). (2012). *Frequency effects in representation*. Berlin: Mouton de Gruyter.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24, 143–188.

Haspelmath, M. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19, 1–33.

Hay, J. (2001). Lexical frequency in morphology: Is everything relative? *Linguistics* 39, 1041–1070.

Hilpert, M., & Diessel, H. (2016). Entrenchment in construction grammar. In Hans-Jörg Schmid (Ed.), *Entrenchment, memory and automaticity. The psychology of linguistic knowledge and language learning* (pp. 45–59). Berlin: Mouton de Gruyter.

Hopper, P. J. (1987). Emergent grammar. *Berkeley Linguistics Society* 13, 139–157.

Pluymaekers, M., Ernestus, M., & Baayan, H. R. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica* 62, 146–159.

Szmrecsanyi, B. (2006). *Morphosyntactic persistence in spoken English. A corpus study*. Berlin: Mouton de Gruyter.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language* 35, 566–585.

Wulff, S. (2008). *Rethinking idiomaticity. A usage-based approach*. London: Continuum.

## References

Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgments of argument-structure overgeneralization errors. *Cognition* 106, 87–129.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62, 67–82.

Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science* 21, 170–177.

Baayen, H. (1993). On frequency, transparency and productivity. In G. Booij & J. V. Marle (Eds.), *Yearbook of morphology* (pp. 181–208). Dordrecht, Netherlands: Springer.

- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3-73). Cambridge, U.K.: Cambridge University Press.
- Bell, A., Brenier, J., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60(1), 92-111.
- Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America* 113(2), 1001-1024.
- Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and development of language* (pp. 279-352). New York, NY: Wiley.
- Boas, H. (2003). *A constructional approach to resultatives*. Stanford, CA: CSLI Publications.
- Bod, R. (2009). From exemplar to grammar: A probabilistic analogy-based model of language learning. *Cognitive Science* 33, 752-793.
- Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32, 45-86.
- Bowerman, M. (1988). The “no negative evidence” problem: How do children avoid constructing an overly general grammar? In J. Hawkins (Ed.), *Explaining language universals* (pp. 73-101). Oxford: Basil Blackwell.
- Bresnan, J., Cueni, A., Nikitina, T., & Baayen, H. R. (2007). Predicting the dative alternation. In G. Boume, I. Kraemer, & J. Zwarts (Eds.), *Cognitive foundations of interpretation* (pp. 69-94). Amsterdam: Royal Netherlands Academy of Science.
- Brooks, P. J., Tomasello, M., Dodson, K., & Lewis, L. B. (1999). Young children’s overgeneralizations with fixed transitivity verbs. *Child Development* 70, 1325-1337.
- Bybee, J. (1985). *Morphology*. Amsterdam: John Benjamins.
- Bybee, J. (1995). Regular morphology and the lexicon. *Language And Cognitive Processes* 10, 425-455.
- Bybee, J. (2001). *Phonology and language use*. Cambridge, U.K.: Cambridge University Press.

- Bybee, J. (2002). Sequentiality as the basis of constituent structure. In T. Givón & B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 109–132). Amsterdam: John Benjamins.
- Bybee, J. (2003). Mechanisms of change in grammaticalization: The role of frequency. In R. Janda and B. Joseph (Eds.), *Handbook of historical linguistics* (pp. 602–623). Oxford: Blackwell.
- Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language* 82, 711–733.
- Bybee, J. (2010). *Language, usage, and cognition*. Cambridge, U.K.: Cambridge University Press.
- Bybee, J., & Hopper, P. (2001) (Eds.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.
- Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37, 575–596.
- Bybee, J., & Slobin, D. (1982). Rules and schemas in the development and use of the English past tense. *Language* 58, 265–289.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Croft, W. (2003). *Typology and universals*, 2d ed. Cambridge, U.K.: Cambridge University Press.
- Dąbrowska, E. (2012). Different speakers, different grammars. *Linguistic Approaches to Bilingualism* 2, 219–253.
- De Smet, H. (2012). The course of actualization. *Language* 88, 601–633.
- Diessel, H. (2005). Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43, 449–470.
- Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology* 25, 108–127.
- Diessel, H. (2008). Iconicity of sequence: A corpus-based analysis of the positioning of temporal adverbial clauses in English. *Cognitive Linguistics* 19, 457–482.

- Diessel, H. (2012a). Buehler's two-field theory of pointing and naming and the deictic origins of grammatical morphemes. In T. Breban, L. Brems, K. Davidse, & T. Mortelmans (Eds.), *New perspectives on grammaticalization: Theoretical understanding and empirical description* (pp. 35–48). Amsterdam: John Benjamins.
- Diessel, H. (2012b). Language change and language acquisition. In A. Bergs & L. Brinton (Eds.), *Historical linguistics of English: An international handbook*, Vol. 2 (pp. 1599–1613). Berlin: Mouton de Gruyter.
- Diessel, H. (2015). Usage-based construction grammar. In E. Dąbrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 295–321). Berlin: Mouton de Gruyter.
- Diessel, H. (2016). Frequency and lexical specificity. A critical review. In H. Behrens & S. Pfänder (Eds.), *Experience Counts: Frequency effects in language* (pp. 209–237). Berlin: Mouton de Gruyter.
- Ellis, Nick C. (1996). Sequencing in SLA. Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition* 18, 91–126.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness. A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text* 20, 29–62.
- Fillmore, C. J., Kay, P., & O'Connor, C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64, 501–538.
- Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language* 66, 789–806.
- Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language* 7, 58–93.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning* (pp. 199–241). London: Cambridge University Press.
- Gerken, L. A. (2004). Nine-month-olds extract structural principles required for natural language. *Cognition* 93, B89–B96.

Givón, T. (1971). Historical syntax and synchronic morphology: An archeologist's field trip. *Chicago Linguistic Society* 7, 384–415.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.

Gómez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by 1-year-olds leads to specific abstract knowledge. *Cognition* 70, 109–135.

Greenberg, J. H. (1966). *Language universals: With special reference to feature hierarchies*. The Hague: Mouton.

Gries, S., & Stefanowitsch, A. (2004). Extending collexeme analysis. *International Journal of Corpus Linguistics* 9, 97–129.

Haspelmath, M. (2008). Creating economical morphosyntactic patterns in language change. In J. Good (Ed.), *Language universals and language change* (pp. 185–214). Oxford: Oxford University Press.

Hilpert, M. (2013). *Constructional change in English: Developments in allomorphy, word-formation and syntax*. Cambridge, U.K.: Cambridge University Press.

Hilpert, M. (2014). *Construction grammar and its application to English*. Edinburgh: Edinburgh University Press.

Israel, M. (1996). The way constructions grow. In A. Goldberg (Ed.), *conceptual structure, discourse and language* (pp. 217–230). Stanford, CA: CSLI Publications.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20, 137–194.

Jurafsky, D., Bell, A., Gregory, G., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: John Benjamins.

Jurafsky, D., & Martin, J. H. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice Hall.

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

- Krug, M. (1998). String frequency. A cognitive motivating factor in coalescence, language processing, and linguistic change. *Journal of English Linguistics* 26, 286–320.
- Krug, M. (2003). Frequency as a determinant of grammatical variation and change. In G. Rohdenburg & B. Mondorf (Eds.), *Determinants of grammatical variation in English* (pp. 7–67). Berlin: Mouton de Gruyter.
- Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Langacker, R. W. (1987). *Foundations of cognitive grammar. Vol. 2. Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 24–63). Stanford, CA: Stanford University Press.
- Langacker, R. W. (2008). *Cognitive grammar. A basic introduction*. Oxford: Oxford University Press.
- Logan, G. D. (1988). Towards an instance theory of automatization. *Psychological Review* 95, 492–527.
- MacDonald, Maryellen C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes* 9, 157–201.
- Manning, C. D. & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science* 283, 77–80.
- Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in the speech to young children. *Cognitive Science* 26, 393–424.
- Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorization. *Cognition* 96, 143–182.
- Newmeyer, F. (2003). Grammar is grammar and usage is usage. *Language* 79, 682–707.

Nosofsky, R. M. (1988). Similarity, frequency and category representation. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14, 54–65.

Pickering, M. J., & Ferreira, V. S. 2008. Structural priming: A critical review. *Psychological Bulletin* 134, 427–459.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46, 115–154.

Pinker, S. (1989). Learnability and cognition. The Acquisition of argument structure. Cambridge, MA: MIT Press.

Pollack, I., & Pickett, J. M. (1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior* 3, 79–84.

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science* 22, 435–469.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.

Schneider, W., & Chein, J. M. (2003). Controlled and automatic processing: From mechanisms to biology. *Cognitive Science* 27, 525–559.

Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science* 23, 569–588.

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language* 81, 105–120.

Spivey-Knowlton, M. J., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition* 55, 227–267.

Stefanowitsch, A. (2008). Negative entrenchment: A usage-based approach to negative evidence. *Cognitive Linguistics* 19, 513–531.

Taylor, J. R. (2012). *The mental lexicon. How language is represented in the mind*. Oxford: Oxford University Press.

Thompson, S. A., & Hopper, P. (2001). Transitivity, clause structure, and argument structure: Evidence from conversation. In Bybee, J. L., & Hopper, P. J. (Eds.), *Frequency and the emergence of linguistic structure* (pp. 27–60). Amsterdam: J. Benjamins.

Tiersma, P. (1982). Local and general markedness. *Language* 59, 832–849.

Tomasello, M. (2003). *Constructing a language. A usage-based theory of first language acquisition*. Cambridge, MA: Harvard University Press.

Tottie, G. (1991). Lexical diffusion in syntactic change: Frequency as a determinant of linguistic conservatism in the development of negation in English. In D. Kastovsky (Ed.), *Historical English Syntax* (pp. 439–467). Berlin: Mouton de Gruyter.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes*. Oxford: Oxford University Press.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language* 35, 566–585.

Wiechmann, D. (2008). Sense-contingent lexical preferences and early parsing decisions: Corpus-evidence from local NP/S-ambiguities. *Cognitive Linguistics* 19, 447–464.

Wiechmann, D., & Kerz, E. (2013). The positioning of concessive adverbial clauses in English: Assessing the importance of discourse-pragmatic and processing-based constraints. *English Language and Linguistics* 17, 1–23.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, U.K.: Cambridge University Press.

Yang, C. D. (2004). Universal grammar, statistics or both? *Trends in Cognitive Science* 8, 451–456.

Zacks, R. T., & Hasher, L. (2002). Frequency processing: A twenty-five year perspective. In Sedlmeier, P., & Betsch, T. (Eds.), *ETC. Frequency processing and cognition* (pp. 21–36). Oxford: Oxford University Press.

**Holger Diessel**

Department of English, University of Jena

**Martin Hilpert**

Institut de langue et littérature anglaises, Université de Neuchâtel

