

# Introduction to frequency and the emergence of linguistic structure

JOAN BYBEE and PAUL HOPPER

## 1. Introduction

A legacy of the structural tradition in linguistics is the widespread acceptance of the premise that language structure is independent of language use. This premise is codified in a variety of theoretical distinctions, such as *langue* and *parole* (Saussure 1916) and competence and performance (Chomsky 1965). A further premise of this legacy is that the study of structure is a higher calling than the study of usage and is a potentially more promising avenue for uncovering the basic cognitive mechanisms that make human language possible.

In contrast, outside linguistics it is widely held that cognitive representations are highly affected by experience. In humans and non-humans detailed tracking of probabilities leads to behavior that promotes survival (Kelly and Martin 1994). Even within linguistics, certain usage-based effects permeate the general lore that practitioners and theoreticians accept: unmarked members of categories are more frequent than marked members (Greenberg 1966); irregular morphological formations with high frequency are less likely to regularize; regular patterns have a wider range of applicability; and high frequency phrases undergo special reduction. Many of these effects had been catalogued and described by George K. Zipf in a pioneering work from the 1930s, *The Psycho-Biology of Language* (Zipf 1965 [1935]). Zipf is known these days chiefly for his "law" that the length of a word is inversely proportional to its frequency and his explanation through the "principle of least effort." While this aspect of Zipf's work is often criticized (see, for example, Miller 1965), he anticipated many of the themes of more recent investigations of the relationship between frequency and structure, such as the fusion of pronouns with auxiliaries in forms like *we're*, *you'll*, etc. and their significance for the genesis of inflection (Zipf 1965 [1935], 247–51). Zipf coined the term "dynamic philology" for the quantitative study of language change and its relevance for linguistic structure.

Zipf's work in linguistics was taken up only sporadically in the discipline as

linguists focused their attention on the theoretical questions of how to define the structural units of language exclusively through local combinatorial possibilities. However, by the 1980s, a number of linguists had begun to think of linguistic structure (grammar) as a response to discourse needs, and to consider seriously the hypothesis that grammar comes about through the repeated adaptation of forms to live discourse (Hopper 1979; Givón 1979; Givón (ed.) 1983; Hopper and Thompson 1980, 1984; Du Bois 1985). The parallel question of how experience with language as reflected in frequency could affect cognitive representations and categorization and thus the internalized grammar of language users also began to occupy researchers at this time in both linguistics (Bybee 1985) and psychology (Rumelhart and McClelland 1986). A related development, symptomatic of the increasing impatience with studies of individual "competence" and growing suspicion regarding the reliability of intuitions as a source of data, was the rise in the 1990s of the new field of corpus linguistics. Starting from trends that had begun with "computational linguistics" going back as far as the 1950s, corpus linguistics has been made possible by the exponential increase in data storage and high-speed processing. While the corpus is the prime tool for frequency studies in general, with many linguists it also serves as a heuristic for new facts about linguistic structure. One especially important claim coming out of corpus studies is that the dividing line between grammar and lexicon, which has been virtually a dogma in linguistics, cannot be sustained (see Stubbs 1996: 36–9 for discussion from the perspective of the Hallidayan strain in corpus linguistics, also Langacker 1987, Hopper 1987, Poplack, this volume, Bybee, this volume, Hallan, this volume, and others). Time and again the operation of linguistic rules has been found to be limited by lexical constraints, sometimes to the point where a construction is valid only for one or two specific words.

Increasingly, then, in many quarters structure has come to be seen not as a holistic autonomous system but as something more fluid and shifting. An influential concept here has been that of *emergence* (Hopper 1987, 1998, 1988, 1993), understood as an ongoing process of *structuration* (Giddens 1984). Structuration in recent sociology refers to "the conditions which govern the continuity and dissolution of structures or types of structures" (Giddens 1977: 120). Emergence in this sense is distinct from *ontogenesis*, which refers to the origins and development—the history—of an existent organism or of a system. By contrast, emergent structures are unstable and are manifested stochastically. The fixing of linguistic groups of all kinds as recognizably structural units (word and phrase units) is an ongoing process; it is the result at any point in time of the "constant resystematization" of language (Coseriu 1954). From this perspective, mental representations are seen as provisional and temporary states of affairs that are sensitive, and constantly adapting themselves, to usage. "Grammar" itself and associated theoretical postulates like "syntax" and "phonology" have no autonomous existence beyond local storage and

real-time processing (Hopper 1987; Bybee, this volume). The notion of language as a monolithic system has had to give way to that of a language as a massive collection of heterogeneous *constructions*, each with affinities to different contexts and in constant structural adaptation to usage (Langacker 1987).

The notion of emergence constitutes a break with standard ideas about grammar that envisage it as a fixed synchronic system. It relativizes structure to speakers' actual experience with language, and sees structure as an on-going response to the pressure of discourse rather than as a pre-existent matrix (Hopper 1988; Ochs, Schegloff, and Thompson 1997). It follows that accounts of grammatical (and phonological) structure must take note of how frequency and repetition affect and, ultimately, bring about form in language (Bybee 1985, to appear; Bybee *et al.* 1994). Now work on the notion that frequency of exposure and use is an important factor in the establishment and maintenance of linguistic structure has begun to branch out in many directions. One of the goals of this book is to represent some of the findings of this research.

### 1.2 Contents of the volume

The papers in this volume build on these two strands of research into language use—the heuristic of frequency and the metalinguistic principle of emergence—to illustrate certain general principles that are robustly documented by empirical investigations of various sorts: distribution in natural conversation, diachronic change, variability, child language acquisition, and experimentation. Two major principles are addressed here:

1. The distribution and frequency of the units of language are governed by the content of people's interactions, which consist of a preponderance of subjective, evaluative statements, dominated by the use of pronouns, copulas and intransitive clauses.
2. The frequency with which certain items and strings of items are used has a profound influence on the way language is broken up into chunks in memory storage, the way such chunks are related to other stored material and the ease with which they are accessed.

Each of the chapters of this volume treats several issues related to these two principles, so that organizing them thematically has been difficult. The organization we have settled is:

- I. **Patterns of Use.** These are papers that deal with patterns of occurrence of morphosyntactic structures in natural conversation;
- II. **Word-level frequency effects,** that is, the papers that deal with the direct and indirect effects of frequency of use on change and structure at the word level;

- III. Phrases and constructions**, which contains papers that demonstrate that many of the same principles found at the word level also operate in multi-word sequences; and
- IV. General**. In this category are placed papers that reference and model multiple phenomena and therefore do not fit easily into the first three categories.

## 2. Patterns of use in natural discourse

### 2.1 Use of natural discourse data

By definition, any study that deals with tokens (as opposed to types) takes as its data base extended samples of natural language, whether these be written language or transcriptions of speech. It might seem that discourse data are simply an extension of the data from intuition, differing only or primarily in quantity but otherwise consistent in structure with forms retrieved through introspection. However, a number of authors in the present volume have drawn attention to what Scheibman (this volume) terms the “slippage” between standard ideas about grammaticality and the facts presented by natural data. Poplack (this volume) finds that the grammar of the subjunctive and conditional in the spoken vernacular French of Canada is quite different from that of the norms dictated for Metropolitan French by the Académie Française. On the other hand, the official grammar fails to note the significant role of lexical constraints in these constructions in the vernacular, where a small number of verbs and main/subordinate verb combinations decisively dominate the grammatical picture. In fact, virtually none of the meanings or functions attributed to the Future or the Subjunctive in official French appear to any significant degree in Canadian French. She echoes a conclusion that has been reached by numerous students of vernacular French (Bauche 1928; Bally 1966 [1932]; Ashby 1977; Lambrecht 1981; etc.), namely that modern spoken French has moved radically beyond the official written language that still forms the basis for structural grammatical studies.

Poplack echoes a theme in frequency studies that is repeated in several of the papers in this volume: there is a very serious mismatch between the results of quantitative studies and grammatical accounts—both descriptive and normative—that rely exclusively on imaginary data. Hallan, commenting on the disparity between the standard view of the prepositional phrase consisting of PREPOSITION + NOUN as the normal context for the category Preposition in English and the markedly late appearance of this pattern in child language, concludes that this prototype is engendered in the course of schooling for literacy. Hallan goes even further in suggesting that the availability of large corpora might call for a general reassessment of grammatical categories. In this she follows Sinclair’s assertion that “even major parts of

speech are not as solidly founded as they might be” (Sinclair:1992: 14, cited in Stubbs 1996: 39).

Hopper and Thompson also note some fundamental differences between linguistic structure as it is posited on the basis of imagined configurations and that of utterances in live conversational contexts. They note that lexical frames for verbs that specify their possible argument structures in advance of usage are often violated in practice, and that the more frequent a verb type, the less predictable the number of arguments; a rare verb like *to elapse* is limited to a single argument, whereas a common verb like *to get* appears in discourse with one, two, or three of the traditional arguments depending on the speaker’s need. Scheibman, arguing for the centrality of subjective expression in conversational English, points out that this role of subjectivity is in opposition to the privileging of referential language in standard linguistic analysis.

### 2.2 Subjectivity

By their nature, all frequency studies are based on usage in some measure. Scholars differ, however, in the degree to which discourse figures as a central part of the study rather than as the site for statistical studies. For a number of contributors, especially those concerned with morphological and phonological questions, the interpretation of contextual meaning is largely irrelevant, since what is at issue for them is type or transitional frequency. For others, it is essential that quantitative work should be combined with a more or less close reading of textual data. Scheibman, for example, shows in detail how the personal, expressive nature of spoken discourse manifests itself in what would be, from the point of view of canonical linguistics, skewed distributions of pronouns and tenses. She stresses that common categories such as third person singular are frequent not because discourse is naturally referential (quite the contrary), but because this category conflates several subtypes, usually evaluative, such as *it in it isn’t fair*. She shows through careful text counts that the relatively high frequency of first and second person singular pronouns is owed to their collocation with verbs of cognition (*I think*, etc.) She concludes that interactive discourse favors “those subject-predicate combinations that permit speakers to personalize their contributions . . .” Further evidence of the frequency of subjectifying elements is the high frequency of modal verbs in the corpus examined in Krug’s contribution to the volume. Modal verbs provide the speaker’s evaluation or perspective on the situation described by the main verb (Scheibman 2000).

Poplack adds another dimension to this same theme, that of Variation Theory, which seeks to identify the different contexts that give rise to the choice of one or another variant of a form. She shows that these contexts can be quite elaborate, and

include a strong element of lexical preference and a less clear influence of type and token frequency. Poplack's domain of study is that of the Irrealis in Canadian French. She shows that robust collocations between the main-clause verb and the verb of the subordinate clause have persisted over long periods of time, and she thus provides historical proof of storage units that transgress the clause boundary.

Thompson and Hopper, in their study of transitivity in spoken discourse, point to incompatibilities between standard accounts of clause structure based on intuitions and the less rigid structure of utterances in conversational contexts. They show that high transitivity in the sense of Hopper and Thompson 1980, which is often taken as the prototype for fully exemplified argument structure, is rare or absent in normal utterances. More generally, the argument structure frames for verbs predicted by theories of the "mental lexicon" are only recoverable for natural discourse to the degree that the verb is unusual. For the more common verbs such as *to get* exceptions and special uses abound to the point of invalidating a priori schemata.

Hallan tackles the problem of "prepositions" or "path morphemes" and their ontogenesis. Tracing the development of such forms as *on* and *over* from their earliest attestations in infants' utterances, she challenges the standard ideas about "prepositions" and the cognitive models of the preposition that are based on a prototype of prenominal forms with locative meanings.

A diachronic perspective on frequency effects is presented by Smith in his study of the English anterior aspect. Smith looks at the distribution of *to be* and *to have* as auxiliaries of the anterior aspect (*I am gone* vs. *I have gone*) in the earliest Old English texts, and makes the point that earlier attempts at synchronic semantic analyses of the distribution have not worked because the synchronic distribution represents a system in flux, caught up in the beginnings of a process whereby the *have* forms are bit by bit encroaching on the *be* forms. Smith hypothesizes that type frequency is a better predictor of the eventual victor in a competition between two forms than token frequency, based on the role of type frequency in morphological productivity (MacWhinney 1978; Bybee 1985). Although the textual frequency of the two auxiliaries is about equal in Old English, the number of verbs construed with *have* (i.e., the type frequency) is by far preponderant. On the other hand, he suggests, high frequency tokens of the less frequent competing types will be the last to succumb to the specialization process. (Poplack, however, finds that for two of her three variables, the future and the imperfect, productivity is not robustly predicted by either type frequency or token frequency, and speculates that other factors may be at work.)

Berkenfield and Bush adopt a more "micro" perspective on discourse in their studies of the "morphophonetics" of English *that* and cross-word boundary palatalization respectively. Berkenfield examines the descendants of the Old English demonstrative pronoun *that* in its functions as demonstrative pronoun, as demon-

strative adjective, as complementizer, and as relativizer from the point of view of the relationship between vowel quality and frequency. She shows strikingly that vowel reduction, as measured by both vowel length (in microseconds) and quality (F1 value), and token frequency go hand in hand with increasing grammaticization. She concludes that "sub-phonemic" distinctions are nonetheless available to speakers in discriminating morphosyntactic functions, a result of some significance for our notions of phonology and phonetics (see also Bybee [this volume, and to appear]).

In another paper influenced by Bybee's "Usage-based Phonology," Bush studies the palatalization of segments across word boundaries in, for example, "would you" > [wudju] as opposed to the absence of such palatalization in sequences such as "good you" (which had been noted by earlier researchers). Bush invokes *transitional probability*, the degree of likelihood that one word will be followed by a specific collocate. He concludes that the discourse "chunking" of lexical words creates units that may behave in every respect like unitary words, permitting the application of processes that are otherwise word-internal (see Bybee 2000a). His study indicates that frequency of cooccurrence significantly drives assimilation whether words are function or content words. Palatalization in conversation is not restricted to the pronoun *you* as suggested by some studies, nor is it possible to predict its occurrence with reference to constituent structure. Pairs of words that are frequently used together, whatever their apparent constituency and status as lexical or grammatical (*don't you, told you, that you, last year*), are more likely to show effects of coarticulation than words that are used together less often.

In most or all of these studies the speech act and its participants have a central role. Most utterances are evaluative in the sense of either expressing a judgement or presenting the world from the perspective of the self or on interlocutor. Referential utterances of the kind that figure so prominently in "syntactic" studies, with their lexical nouns and single-word verbs, are in fact rare in natural discourse. They favor third-person constative utterances, ones that typically indicate a shift of the focus of the discourse away from its immediate existential context and into the realm of unwitnessed, objective, remote events (Hopper 1997). Natural discourse is concerned with the here-and-now world of the speaker and the hearer, and with the contingencies (imperatives, conditionals, possible worlds) that proceed directly from it. Natural discourse is, in other words, preeminently *subjective* (Scheibman, this volume, and Scheibman 2000). If grammar is emergent from commonly used sequences, it is natural to expect that that such sequences will comprise the core of grammaticalized structures, and therefore that grammar—the internalized aggregate of formations from usage—will move into increasingly subjective spheres (Traugott 1989; Traugott and Koenig 1991). This point is implicit or explicit in a number of papers in which functional areas like modality, transitivity, and aspect figure.

### 3. Units of usage

A mainstay of linguistic analysis is the identification of recurring units in the continuous stream of speech. Thus we identify features, segments, syllables, morphemes, words, phrases, clauses and so on, on the basis of distribution and with reference to phonetic and semantic features. The problems with attempting to exhaustively and discretely divide utterances into these units are well-known: classificational difficulties include the issues of diphthongs, affricates, extra-syllabicity, empty morphs, clitics, auxiliary verbs, and subordinate clauses, yet the reality of these recurring units is attested to by their patterns of use, including use in novel utterances. An important motivation for identifying these units is to assign them to lexical storage or to describe their structure in terms of grammatical rules. It has been assumed that language users come up with the same analysis as linguists do and that the most economical treatment of this complex system would postulate a small number of storage units and a set of rules for their combination. Thus it is usually assumed that morphemes or words are units of storage and access, while larger units are produced by combinatory rules.

Recently evidence from a variety of research traditions has been brought forward to question the economy of storage and the separation of lexicon from grammar (Hopper 1987; Langacker 1987; Stubbs 1996; etc.). The model that is emerging to replace the old conception postulates that to a large extent, the units of usage ARE the units of storage and access. As people do not speak in isolated morphemes or words, in many cases the units of memory and processing contain multiple morphemes and even multiple words (see Wray and Perkins 2000). The categorization in storage of units of use forms a network based on the user's experience with language, and from this network, recurrent patterns emerge (Bybee 1998).

What sorts of units might we expect to find in storage? We are largely restricted to answering this question for English, since so much more research has been concentrated on this language than on any others. First, the traditional unit of noun phrase does occur in conversation, and NPs are often independent intonation units (Ono and Thompson 1994; Croft 1995). NPs almost always include determiners (*alan* or *the*) and the phonological alternations that are characteristics of these determiners suggest that DET + NOUN is a storage unit (Bybee, to appear).

Verbs in most languages are multi-morphemic units (given the widespread occurrence of inflection on verbs), but in English verbal expressions are typically dispersed over multiple words. Hopper 1991 cites examples of VERB + PARTICLE, VERB + ADVERB, VERB + PREPOSITION, VERB + NOUN, AUX + VERB in which the VERB element is not readily separable from other parts of the functional group. Examples are *wake up*, *speeded up*, *head straight in*, *has drifted left*, *heave a sigh of relief*, *start exploding*, *have to quickly decide* and so on.

Other sorts of frequently-occurring snatches of speech that show evidence of autonomy characteristic of stored items are *I don't know*, *I don't think* (Bybee and Scheibman 1999), *wanna*, *gotta*, *gonna*, etc. (Krug, this volume), *you and I* (Boyland, this volume), *did you*, *didn't you*, *don't you* (Bush, this volume), *is gone* (Smith, this volume), *come on*, *over here*, *over there* (Hallan, this volume). Stubbs (1996: 39), citing Renouf and Sinclair (1991), adds from corpus studies *a couple of*, *a lot of*, *an indication of*, *an element of*, *be able to*, *be ready to*, *too easy to*, *too close to*, concerning which he comments: "Such items are highly frequent, an integral part of language, yet lie somewhere between word and group." These frequently-occurring sequences include both phrases that can be used in isolation, e.g. *I don't know*, and also parts of constructions that require nouns and verbs to be combined with them to be complete: NP *wanna* VERB, NP *be able to* VERB, NP *gonna* VERB, *did you* VERB, NP *is gone*. Thus many of the storage units we are proposing here are constructions in that they have open slots that take items that share certain properties.

In the model suggested by the papers in this volume (Pierrehumbert, this volume; see also Bybee 1998, to appear) tokens of experience with language are organized into exemplars on the basis of high similarity of phonetic shape and function or meaning, and such exemplars are tagged for their contextual associations, both linguistic and extra-linguistic. Thus tokens of *I* in *I don't know*, *I don't think*, *I see*, *I want* etc. are mapped onto the same representations. This does not prevent a strong link between *I* and *don't* from also being maintained, as *don't* is the second most frequent item to follow *I* (*'m* is the most frequent) (Bybee and Scheibman 1999; Krug 1998). Thus even though complex units (such as *I don't*) are stored and accessed, their component parts are also identified in the categorization and storage process.

A major part of the evidence for the storage of multimorphemic words and multiword phrases and constructions is the fact that, as shown in several of the papers of this volume (see especially, Part 2), both direct and indirect frequency effects can be demonstrated for these units. Linguistic material cannot accrue frequency effects unless the brain is keeping track of frequency in some way; frequency effects cannot be attributed to units unless they are items in storage that are affected by experience. A natural way to track frequency is to postulate that tokens of experience strengthen stored exemplars (Bybee 1985; Pierrehumbert, this volume).

In the following sections, we will discuss the effects of frequency that have been documented in the literature and in the papers of this volume as applying both at the word level and in multi-word sequences and the cognitive mechanisms underlying them.

#### 4. Frequency effects and cognitive mechanisms in emergent grammar

The notion of emergent structure has become important in various branches of the sciences in the last two decades. The basic idea is that what may appear to be a coherent structure created according to some underlying design may in fact be the result of multiple applications or interactions of simple mechanisms that operate according to local principles and create the seemingly well-planned structure as a consequence. MacWhinney (this volume) discusses emergentist theories as they have developed in the physical and biological sciences, and examines the various ways in which emergence can be applied to the study of language. His discussion focuses on neurological models of language learning and representation that proceed from local self-organizing maps to more complex networks that incorporate larger chunks of language, multiple associations among the units of language as well as grounding in the physical and social domains. Such models can accommodate frequency effects, as long as the details about how frequency effects work can be established empirically.

Many of the papers in the current volume are directed towards understanding the multiple ways that frequency of use can effect linguistic behavior. In the following subsections we will discuss these frequency effects focusing on the cognitive mechanisms that bring them about and functional consequences they have for language. These effects are (1) phonological reduction in high frequency words and phrases (4.1); (2) functional change due to high frequency (4.2); (3) frequency and the formation of constructions (4.3); (4) frequency and accessibility (4.4); (5) the retention of conservative characteristics (4.5); and (6) the notion that a stochastic grammar is a result of linguistic knowledge based on experience (4.6).

##### 4.1 *Phonological reduction in high frequency words and strings*

Recent research has documented a tendency identified in Schuchardt 1885 by which words of higher frequency tend to undergo sound change at a faster rate than words of lower frequency. This effect has been identified for English reduction to schwa (Fidelholtz 1975), schwa deletion (Hooper 1976), t/d deletion (Bybee 2000, Gregory *et al.* 1999), deletion of [ð] in Spanish (Bybee, to appear), the raising of /a/ before nasals in Old English (Phillips 1980), and the raising of /æ/ in San Francisco English (Moonwomon 1992).<sup>1</sup> Among the current papers, the effects of high frequency on reductive change is documented in the chapters by Berkenfield, Bush, Jurafsky *et al.*, Krug, and Phillips; it is also discussed in the chapter by Fenk-Oczlon.

One of the most important consequences of these studies is the finding that sound change is gradual both phonetically and lexically, because this means that very specific phonetic features, probably specified as a range of phonetic variation, are

associated with particular lexical items (Hooper 1981; Bybee 2000b, Mowrey and Pagliuca 1995). Any phonological representation that fails to register non-contrastive features is not able to account for this lexically-specific variation. On the other hand, an exemplar model, as proposed in Johnson 1997, which records and organizes in memory distinct phonetic variants of words and phrases, can accommodate lexical variation.

Pierrehumbert (this volume) demonstrates how an exemplar model can be formalized to account for both the perception and production of lexically-specific variation. In addition, she models the effect that a lenition bias or tendency towards reduction can have on a set of exemplars, and the effect of token frequency on reductive processes.

The origins of reduction are in the automatization of neuro-motor sequences which comes about with repetition. This automatization involves the reduction of the magnitude of articulatory gestures and the increased overlap of these gestures (Browman and Goldstein 1992; Mowrey and Pagliuca 1995). Such reductions are systematic across speakers; that is, they do not represent 'sloppy' or 'lazy' speech. Moreover, reduction or lack of reduction are carefully monitored and controlled by the speaker according to the context. As a result, reduction or lack of it can take on pragmatic value.

The role of token frequency in reductive sound change involves the interaction of a complex set of factors. One factor is that automatization is occurring whenever speech is produced, which results in small changes in the magnitude and timing of gestures; frequent words have more opportunity to be affected as they are exposed to these on-line processes more than infrequent words (Moonwomon 1992). Frequent words are also used more in familiar, casual settings, where more reduction is allowed than in formal settings. This also exposes frequent words to more reduction. This point is also made by Dahl (this volume), who points out that a number used as a date, such as "1999", characteristically receives a more reduced articulation than the same number used to denote a quantity or a street address.

In addition, Fowler and Housum (1987) have shown that in reading a narrative, subjects' productions of the second occurrence of a word in the narrative is significantly shorter in overall duration than the first occurrence of the word. The shortening of a word has an effect on all the gestures that comprise the word, decreasing their magnitude and increasing their overlap. Gregory *et al.* (1999) find a similar effect in conversation, and in addition, report the semantic relatedness of a word in the discourse has a very strong effect on duration. Thus, when the word *coast* occurred in a conversation about weather it was much shorter than when it occurred in a conversation about family budgets.

The speaker seems to be able to gauge how much phonetic information the hearer needs in order to access the correct word. Where the word has occurred before, it

is primed and easier to access; where the word is primed by the other words in the context, it is also easier to access. The persistent use of this strategy by speakers leads to the development of a listener strategy by which reduced words are judged to be repetitions and thus part of the background in the discourse (Fenk-Oczlon, this volume). Thus with the reduction the speaker signals that the reduced word is just the same old word as used before, not a new one.

Fenk-Oczlon (this volume) relates these correspondences to information flow: according to her, efficiency demands a relatively constant flow of information. Thus short words should convey less information than longer words. Besides relating to the correspondence between length of word and semantic complexity, this principle makes several predictions about the length of words and their position in discourse ('more frequent before less frequent'), in particular in binomial 'freezes' (frequently conjoined nouns) such as *bread and butter*, *salt and pepper*. By showing a discourse relationship that goes beyond one of mere length-to-frequency, but rather places these in a functional frame, she operationalizes Zipf's "law" and strips it of the standard objection that the law amounts to a tautology (see Miller's introduction to Zipf 1965: v-x).

The paper by Jurafsky *et al.* (this volume) takes into account a number of factors under the Probabilistic Reduction Hypothesis, which includes not just the predictability of a word within a particular discourse, but also its cumulative token frequency and the probability of a word given neighboring words. Jurafsky *et al.* provide useful formulae for calculating the predictability of a word given the previous and following word. They study the top ten most frequent words of English, which are all function words (*a, the, in, of, to, and, that I, it, you*). These words both show more vowel reduction and shorter duration as they are more predictable from the preceding and following word. In contrast, content words ending in /t/ or /d/ were studied for the deletion of their final consonant and here they find that only the frequency of the word containing the /t/ or /d/ predicts the rate of deletion.

Thus Jurafsky *et al.* suggest that function words are more affected by context than the less frequent content words. This in turn indicates that the phonological shapes of function words are more determined by the constructions that they are in, while content words are more independent. In fact, one could argue that function words only occur in constructions and do not have independent representation (unless they are homophonous with a noun or verb). This would mean that function words have multiple representations, since each construction a function word occurs in requires a representation. Berkenfield (this volume) demonstrates that the polysemous function word *that* has different phonetic properties depending upon whether it is functioning as a demonstrative, a complementizer, or a relative clause marker, and that part of this difference is due to the frequency of the different constructions in spoken discourse.

Krug's study (this volume) of the new emerging modals also demonstrates phonological reduction inside frequently occurring chunks: phrases such as *want to, have to, got to, ought to, going to* undergo extraordinary reduction, due, at least in part, to the high frequency with which they are used together. His study also underscores the categorizing features of linguistic storage, as he argues that these units are classed together as a new emergent category of modal auxiliary based on their phonological and semantic coherence.

These studies all lead us to consider the nature of the storage and processing units in mental representation. In order for sequences to accrue frequency effects or phonological fusion and reduction, they must exist as units in mental representation. Jurafsky *et al.* take their evidence to indicate that "probabilistic relations between words are represented in the mind of the speaker" (p. 2). By this they do not mean that any words that affect one another are stored as single units, because they distinguish between lexicalization, by which sequences are treated as single words, and probability relations. A possible interpretation of their findings would be that stored words are linked sequentially and that frequency of co-occurrence strengthens these links.

MacWhinney (this volume), while treating many issues in the modeling of emergence in language, discusses the formation of chunks and advocates a distinction between a chunk in perceptual processing and an avalanche, which is a serial string of behavior, such that the triggering of the beginning of the string leads to the firing of all the component pieces.

#### 4.2 Functional change due to high frequency

Functional and semantic change in high frequency strings or constructions is the focus of the recent research in grammaticization (Bybee *et al.* 1994; Hopper and Traugott 1993, among many others). Phonological reduction and fusion in grammaticization are paralleled by semantic generalization and functional shifts. Frequency is one of the factors that conditions functional change. Haiman (1994) argues that repetition is one of the factors behind emancipation, the process by which an instrumental act becomes symbolic through association with a particular outcome. Repetition also conditions bleaching through the process of habituation, wherein an organism ceases to respond at the same level to a repeated stimulus. Dahl (this volume) likens the process to inflation in economics: as strong expressions are increasingly overused, their effect weakens, and newer, stronger expressions must take their place if the same rhetorical effect is to be achieved.

Grammaticization is the mechanism by which structure emerges from language use. Since such a vast literature on the topic now exists, it is not a specific focus for the current volume. However, the paper here by Krug deals with the emerging class

of modals in English, which are erstwhile main verbs undergoing grammaticization into a new class of modal auxiliaries.

#### 4.3 *Frequency and the formation of constructions*

Constituent structure is determined by frequency of co-occurrence (Bybee and Scheibman 1999): the more often two elements occur in sequence the tighter will be their constituent structure. The tightest constituency is the result of two very specific items occurring frequently together. Clear examples are cases in which two words have fused because of their frequent co-occurrence and now behave essentially as single words, e.g. *want to* > *wanna*, *going to* > *gonna*, *I am* > *I'm*, *can not* > *can't*, *do not* > *don't*, *I don't know* > *I dunno*, *would have* > *would've* (Boyland 1996; Bybee and Scheibman 1999; Krug 1998, this volume). In a sense, frequency of use has led to the loss of former constituent boundaries within these strings, in some cases, major constituent boundaries such as that between subject and predicate (*I'm*) or between main clause and subordinate clause (*wanna*). In addition to these cases, which one might want to view as marginal, the kind of constituency normally studied by syntacticians also has its source in language use and frequency of co-occurrence. Thus determiners occur with nouns, auxiliaries with verbs, prepositions with noun phrases and so on. Constructions such as [DET + NOUN], [AUX + VERB], [PREP + NP] are conventionalized through frequent use.

However, grammars (however conceived) do not merely contain the highly schematic representations such as [DET + NOUN] but also many more specific or local representations with very explicit lexical material included. Again, which representations of this sort exist depends almost entirely on frequency of use. Hallan (this volume) shows that the use of specific instances of constructions begins very early in language acquisition; in fact, children acquire very specific instances of constructions and use them quite appropriately long before there is any evidence for the extraction of grammatical principles from the ambient language. For instance, in the Wells corpus the five most frequent uses by all speakers of the word *over*, often regarded as a canonical preposition, do not include a prepositional use. Instead *over there* occurs the most, followed by *over* in phrasal verbs, such as *fall over*, *knock over* and the phrase *all over (the)*; *over here* is the fifth most frequent use. Similarly the uses of *on* include many fixed expressions and frequent phrases. Not surprisingly, the early uses of *on* are dominated by particular phrasal verb combinations such as *come on*, *put on*, *turn on*. Even the prepositional uses of *on* occur with certain nouns, such as *on the floor*, *on (one's) own*, *on the bed* and so on. This view of child language makes it clear that children acquire very specific expressions and routines that only later become productive and show evidence of more schematic representation (see also Lieven *et al.* 1997).

Boyland's study shows that hypercorrect forms of English pronouns, specifically the use of *you and I* in object position and the use of *whom* in subject position, are heavily influenced by frequency of use. Boyland shows that it is not just any conjoined pronouns that are used in subject form in object position, but rather that it is *you and I* that are most commonly used hypercorrectly. Her corpus study shows moreover that of all conjoined phrases with *I*, the specific phrase *you and I* is the most frequently occurring. She argues that *you and I* has become a processing unit because of its frequency and that, like other high frequency items, it is easy to access whole. Thus speakers use it as a unit rather than generating two separate pronouns with the conjunction. There is of course no question that speakers know that the phrase *you and I* consists of three words, all of which are used elsewhere, but that does not prevent them from packaging this particular sequence as a unit.

Hypercorrect uses of *whom* are also highly influenced by frequency. Local constructions consisting of preposition + *whom* are the most stable and consistently used. Hypercorrection occurs most commonly in larger syntactic units where *whom* plays different roles at different levels, for instance in *Someone whom he feels is worth listening to has convinced him*. Here the smaller clause, *whom he feels*, determines the case of the relative pronoun. The frequency of *whom* use in such cases reflects the frequency of *whom* use in normal relative clauses: it is most frequent after prepositions, next most frequent as a direct object and least frequent as a subject.

Bybee's paper suggests an understanding of French liaison phenomena as a function of constructions in which certain phonological material is highly entrenched. Liaison consonants appear in DETERMINER + NOUN constructions, CLITIC PRONOUN + VERB constructions, prepositional phrases, and some ADJECTIVE + NOUN constructions. Liaison alternations are maintained most consistently in the higher frequency constructions. The fact that frequency plays a greater role than syntactic constituency in determining liaison is brought home by the fact that the third singular of *être* 'to be', *est*, has a much higher rate of maintaining liaison than any other of the forms of the same verb even in the same construction. Bybee proposes that for each construction with the alternation there are two subschemas, one supplying the consonant before a vowel-initial word and one without the consonant before a consonant-initial word. Because consonant-initial words are twice as frequent as vowel-initial words, the latter construction generalizes, bringing about the loss of the liaison consonant. But her study goes further and raises a more general question about the relationship between syntax and phonology. This relationship has generally been seen as a very indirect one between a set of categories and a set of phonological segments. Bybee suggests instead that liaison in French is neither morphosyntactic nor phonological, but is a frequency effect such that "the higher the frequency of a phrase or construction, the more likely it is to preserve liaison." It follows that constituent structure as it is normally viewed, that is, without

reference to the discourse frequency of the set, cannot be causally involved in the loss or preservation of liaison. Speakers, she suggests, CREATE structure by frequent use of certain word combinations, and since each of these words can participate in other collocations, more than one constituent analysis may be possible (for example, 'll in *I'll do it* is simultaneously part of *I* and of *do*.)

#### 4.4 Frequency and Accessibility

It has long been known that the speed of lexical access of individual words is highly affected by frequency of use: in lexical decision tasks, subjects identify words much faster if they are of high frequency. Evidence already reviewed here suggests that frequency of use may make access of larger units easier as well. Strings such as *you and I*, *come on*, *fall over*, and common sequences with liaison in French, such as *mes amis* 'my friends', *c'est un* 'it's a', and *l'un avec l'autre* 'with one another' may be more efficiently accessed as units than composed morpheme by morpheme. In more traditional models in which only monomorphemic units are stored, however, no frequency effects in the processing of multi-morphemic units are to be expected.

Hare *et al.* investigate this issue for regular morphologically complex verbs in English. The hypothesis has been put forward that while irregular English verbs have lexical listing and thus show effects of frequency in derived forms, such as the Past tense, regulars are derived by rule and thus their Past tenses can show no frequency effects (Pinker 1991, among others). Hare *et al.* are able to falsify this hypothesis by demonstrating in two different tasks that subjects respond to regular Past tense words in English according to their frequency of use. Their experiments involve homophones in which one word is a regular Past tense form. The first experiment demonstrates that when subjects are asked to write down the word that they heard there is a strong tendency to write the most frequent of the two homophones, even if that is a regular Past tense form. In the second experiment homophones were used for primes in a lexical decision task. The results showed that if the Past tense member of the homophone pair was more frequent, priming effects on the base verb were evident, but if the non-verb homophone was more frequent, reaction times were slower. These experiments show that regular Past tense verbs in English can accrue lexical strength and thus must have a mental representation.

If regular morphologically complex words can be stored in memory even though they are derivable by regular rules, then there is no reason to suppose that lexicon and grammar are separate from one another (Bybee 1998; Hopper 1987; Langacker 1987; Sinclair 1992). Furthermore, the argument can be taken to a higher level: sequences of words that are frequently used can be represented mentally by the same principle.

#### 4.5 Retention of conservative properties in high frequency units

Hooper 1976 pointed out a paradox in the lexical diffusion of sound change versus analogical change: sound change affects high frequency items first, while analogical change affects low frequency items first. This difference reflects distinct motivations and cognitive mechanisms for the two types of change. High frequency words are in the forefront of phonetic reduction because their frequent use exposes them to reductive processes and their reduction reflects their predictability in discourse (see section 4.1); high frequency forms resist analogical change, such as regularization of irregulars, because their frequency makes them easy to access whole and there is no need to re-form them by regular rule. One of the difficulties in the articulation of this theory of lexical diffusion lies in the specification of the set of changes that proceed from low frequency to high frequency items. In this volume, Phillips proposes that the defining characteristic of such changes is their base in lexical analysis, or the analysis of other forms. Thus when *wept* is replaced by *wepted* it is because of a pattern extracted from the analysis of other forms. Phillips' study contains several examples of word-level changes that proceed in this manner. Since high frequency irregulars are highly entrenched and easily accessible, they are the last to undergo such changes. This pattern of lexical diffusion explains why irregularity is situated in the high frequency paradigms of a language.

To our knowledge, the most comprehensive study of the effects of frequency on the maintenance of morphological irregularity is found in the paper by Corbett *et al.* in this volume. These authors examine and classify all the irregularities found in Russian noun paradigms and count the token frequency of the forms with these irregularities in the one-million-word Uppsala corpus. Their results show a consistently strong association of high token frequency with irregularities in the plural forms of the noun paradigms. They also test the hypothesis that irregularity will be found in plurals which have a high frequency relative to their singulars, but the support for this hypothesis in their data is much weaker. It appears that it is not so much paradigmatic relations as pure token frequency that allows irregularity to be maintained.

For linguistic theory the major consequence of the finding that high frequency units are resistant to reformation on the basis of productive patterns is that the resistant units must have storage in memory in order to resist change and in order to be affected by frequency of use. For irregular nouns and verbs this proposition is not very controversial, since most linguists would now agree that irregular forms have lexical listing. However, the same argument is applicable to syntactic or multi-word units, which also maintain irregular or conservative patterns when of high frequency. The resulting implication that high frequency phrases are stored in memory radically changes our notions of the way syntax operates (Bybee and Thompson 1997).

Smith's study in this volume shows that the maintenance of the *be* auxiliary for the resultative/perfect in English occurred only with the most frequent verbs used in this tense. As the *have* auxiliary generalized to more and more verbs, it worked from the least frequent to the most frequent, leaving only *is gone* in Present Day English. Thus the generalization of the *have* auxiliary worked like a regularization process; in order for the high frequency combinations to resist regularization, phrases such as *is come*, *is gone*, etc. had to have been stored in memory and accessed as units. Bybee's study of French liaison also finds the liaison alternations maintained only in the most frequent constructions, as mentioned above. The loss of liaison is the result of a 'regularization' or the generalization of the construction which does not contain the liaison consonant.

Thus evidence for the storage of auxiliary + verb, determiner + noun, clitic pronoun + verb, and adjective + noun are found in these studies. Poplack's study extends the reach of the storage of specific constructions and phrases even further. Her study of the maintenance of the Subjunctive in Canadian French shows that very specific lexical dependencies can reach across traditional clause boundaries. Only certain high frequency matrix verbs and embedded verbs create the conditions for the appearance of the Subjunctive, suggesting that very specific constructions replete with particular lexical material are accessed to produce the Subjunctive.

#### 4.6 Stochastic grammar

Grammatical generalizations are at their very base variable and probabilistic in nature and derived from the user's experience with language (Pierrehumbert 1994a). Probabilistic knowledge of variation ranges from phonetic detail to word structure to morphosyntactic patterns. Pierrehumbert in this volume demonstrates how phonetic variation can be built into a stochastic grammar.

Frisch *et al.* take up the topic of word structure or phonotactic patterns. Recent research into word and syllable structure has shown that speakers' judgements of acceptability for nonce words corresponds closely to the frequency of those structures in the existing lexicon (Pierrehumbert 1994b; Vitevitch *et al.* 1997). Thus the phonotactic patterns in the lexicon turn out to be good testing ground for the nature of linguistic generalizations. Frisch *et al.* pursue this line of investigation of phonotactic patterns in English and Arabic. They find evidence that subjects use generalizations about existing words at varying levels of abstraction. Their results indicate that subjects used knowledge of natural classes, particularly in making judgements about consonant sequences in Arabic, and that they also used comparison to particular existing words if the nonce word was highly similar to an existing word. Of considerable interest also is their finding for English-speaking subjects, that experience with language changes intuitive judgements: subjects with larger

vocabularies judged low probability words as more acceptable than did subjects with smaller vocabularies. Boyland's study also suggests that speaker's intuition and usage can change over time due to exposure to different patterns.

Poplack (this volume) demonstrates that in some cases variability can be quite stable over long periods of time. Her study indicates that morphosyntactic constructions are not in a simple one-to-one relationship with pragmatic or semantic function, but rather that variant constructions for the same function can alternate freely and this variation can be maintained over a long period of time.

In the view of these authors and others in this volume, grammar is not fixed and absolute with a little variation sprinkled on the top, but it is variable and probabilistic to its very core. Patterns of usage and particular choices made by speakers at any given moment are heavily influenced by both immediate and long range experience with language. Intuitions about grammaticality are based on this experience. An utterance is judged as grammatical if it is highly similar to other frequently heard utterances; if an utterance has a part which bears no resemblance to any previously experienced constructions or fixed phrases, it will be judged to be ungrammatical. Clearly, the criteria for such comparisons with past experience are individual, inexact, and scarcely amenable to treatment in terms of precise objective categories.

## 5. Conclusion

The study of frequency effects in language has important implications for the goals of linguistics. Among other things, it raises the challenging question of how linguists are led to impute structure to any sequence of forms if not on the grounds of their prominence in usage and memory, that is, their usefulness in discourse reflected in their frequency. In other words, what are the alternatives to frequency as an explanation for structure and regularity in language?

One well-known answer to this question involves "intuition," the introspective sense that a sequence conforms to an internalized grammar. But even intuitions could be based on the user's experience with language rather than on an abstract grammar autonomous from language function and use. The dominant paradigms of linguistics assume some such pre-existent holistic grammar as the most important prerequisite for communication. But, as Roy Harris (1990 [1978]: 149) pointed out in his inaugural lecture for the Chair of General Linguistics at Oxford University, such an assumption is suspect on several grounds, in the first place because it is radically at odds with all other forms of social experience. We do not communicate through reference to prior fixed abstract forms, but rather "... we create language as we go, both as individuals and as communities, just as we create our social

structures, and our forms of artistic expression, our moral values, and everything else in the great complex we call civilisation.” Harris calls for “. . . [a movement] away from a study based on the hypothesis of fixed monolithic structures called “languages” which somehow exist independently of whether or how they are brought into use” (Harris 1990 [1978]): 150), and in its place appeals for “. . . a science in which a language is envisaged, not as something which exists as a system over and above the communication situations in which it is manifested, but as a cumulative product of such situations which can be variously exploited to provide a basis for their subsequent renewal.”

In Harris’ “integrational linguistics” (see Toolan 1996 and Wolf ed. 2000 for extended discussion), language would be studied not as a distinct and separately apprehended “segregated” entity but as an activity blended with the nexus of other activities that form part of communicative situations such as chatting or interviewing. On the one hand, the papers in the present volume may be said to constitute a step in this direction. As a body they support Harris’ view of structure as a product rather than a prerequisite of communication. On the other hand, the papers are virtually unanimous in emphasizing the *individual* speaker, focusing on solitary linguistic behavior and cognitive capacities with only fleeting references to the complex of communal experiences which make utterances possible in the first place. But as some of the papers hint, frequency and emergent structure involve more than unmediated linguistic behavior. Situations and their participants are also repetitive phenomena, and linguistic routinization is ultimately inseparable from cultural practices in general. In this respect while these papers retain a link with the standard assumptions of linguistics, they at the same time suggest a basis and a direction for future research.

## Notes

1. It is not the case that sound changes divide neatly into those that evince lexical diffusion and those that apply across the board (the Neogrammarian changes) as proposed by Labov 1981 (see Phillips 1984). Nor is it the case that lexical diffusion of phonetically-motivated change is a type of analogy as claimed by Kiparsky 1995 (see Phillips, 1998).

## References

Ashby, William. 1977. *Clitic inflection in French: A historical perspective*. Amsterdam: Editions Rodopi NV.

- Bally, C. 1966 [1932]. *Linguistique générale et linguistique française*. Fourth Edition. [First published in 1932] Bern: Francke.
- Bauche, Henri. 1928. *Le langage populaire. Grammaire, syntaxe, et dictionnaire du français tel qu'on le parle dans le peuple de Paris. avec tous les termes d'argot usuel*. Paris: Payot.
- Boylard, Joyce Tang. 1996. *Morphosyntactic Change in Progress: A Psycholinguistic Approach*. Dissertation: Linguistics Department, University of California.
- Browman, Catherine P. and Louis M. Goldstein. 1992. Articulatory phonology: an overview. *Phonetica* 49: 155–80.
- Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Philadelphia: Benjamins.
- Bybee, Joan L. 1998. The emergent lexicon. *CLS* 34.2: *The Panels*. 421–35.
- Bybee, Joan. 2000a. Lexicalization of sound change and alternating environments. In *Laboratory V: Language Acquisition and the Lexicon*, Michael Broe and Janet Pierrehumbert (eds.), 250–68. Cambridge: Cambridge University Press.
- Bybee, Joan L. 2000b. The phonology of the lexicon: Evidence from lexical diffusion. In *Usage-Based Models of Language*, M. Barlow and S. Kemmer (eds.), 65–85. Stanford: CSLI.
- Bybee, Joan. To appear. *Phonology and Language Use*. Cambridge University Press.
- Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Bybee, Joan, and Joanne Scheibman. 1999. The effect of usage on degree of constituency: the reduction of *don't* in American English. *Linguistics* 37.575–596.
- Bybee, Joan and Sandra Thompson. 1997. Three frequency effects in syntax. *BLS* 23.65–85.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass: MIT Press.
- Coseriu, Eugenio. 1974. *Synchronie, Diachronie und Geschichte* (trans. Helga Sohre). Munich: Fink. (=International library of general linguistics 3) (*Sinchronía, diachronía y historia; el problema del cambio lingüístico*. Montevideo: Impresora Cordon 1958).
- Croft, William. 1995. Intonation units and grammatical structure. *Linguistics* 33.839–82.
- DuBois, John. 1985. Competing motivations. *Iconicity in syntax*, John Haiman (ed.), 343–65. Amsterdam: John Benjamins.
- Fidelholtz, James L. 1975. Word frequency and vowel reduction in English. *CLS* 11.
- Fowler, Carol A., and Jonathan Housum. 1987. Talkers’ signaling of “new” and “old” words in speech and listeners’ perception and use of the distinction. *Journal of Memory and Language* 26: 489–504.
- Giddens, Anthony. 1977. *Studies in Social and Political Theory*. London: Hutchinson.
- Giddens, Anthony. 1984. *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge: Polity Press.
- Givón, T. 1979. *On Understanding Grammar*. New York: Academic Press.
- Givón, T. 1983. Introduction to Topic Continuity. In *Topic Continuity in Discourse: A Quantitative Cross-Language Study*, T. Givón (ed.). Amsterdam: John Benjamins. (TSL 3).
- Givón, T. (ed) 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam: John Benjamins. (TSL 3)

- Greenberg, Joseph H. 1966. *Language Universals*. The Hague: Mouton.
- Gregory, Michelle, William D. Raymond, Alan Bell, Eric Fosler-Lussier and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. *CLS* 35.
- Haiman, John. 1994. Ritualization and the development of language. In *Perspectives on grammaticalization*, William Pagliuca (ed.), 3–28. Amsterdam: John Benjamins.
- Harris, Roy. 1978. *Communication and Language*. Oxford University Press. Reprinted in *The Foundations of Linguistic Theory. Selected Writings of Roy Harris*, Nigel Love (ed.) 1990, 136–50. London: Routledge.
- Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In *Current Progress in Historical Linguistics*, W. Christie (ed.), 96–105. Amsterdam: North Holland.
- Hooper, Joan Bybee. 1981. The empirical determination of phonological representations. In *The Cognitive Representation of Speech* Terry Myers, John Laver, and John M. Anderson (eds.), 347–57. Amsterdam: North Holland.
- Hopper, Paul J. 1979. Aspect and foregrounding in discourse. In *Discourse and Syntax*, T. Givón (ed.). New York: Academic Press, 213–41.
- Hopper, Paul J. 1987. Emergent grammar. *BLS* 13: 139–57.
- Hopper, Paul J. 1988. Emergent Grammar and the A Priori Grammar Postulate. In *Linguistics in Context*, D. Tannen, (ed.), 117–34. (Collected General Lectures from the 1985 LSA Linguistics Institute, Georgetown University.) Ablex Corp.
- Hopper, Paul J. 1991. Dispersed verbal predicates in vernacular written narrative. *BLS* 17: 402–13.
- Hopper, Paul J. 1997. Discourse and the category Verb in English. *Language and Communication. Special Issue: The Importance of Theory in Discourse Analysis*. Vol. 17.2: 93–102.
- Hopper, Paul J. 1998. Emergent Grammar. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Michael Tomasello (ed.), 155–75. Mahwah, NJ/London: Lawrence Erlbaum Associates.
- Hopper, Paul J. and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56: 251–99.
- Hopper, Paul, and Sandra Thompson. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60: 703–52.
- Johnson, Keith. 1997. Speech perception without speaker normalization. In *Talker variability in speech processing*, Keith Johnson and John W. Mullennix (eds.), 145–65. San Diego: Academic Press.
- Kelly, Michael H. and Susanne Martin. 1994. Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition and language. *Lingua* 92: 105–40.
- Kiparsky, Paul. 1995. The phonological basis of sound change. In *The Handbook of Phonological Theory*, John A. Goldsmith (ed.), 640–70. Oxford: Basil Blackwell.
- Krug, Manfred. 1998. String frequency: a cognitive motivating factor in coalescence, language processing and linguistic change. *Journal of English Linguistics* 26: 286–320.
- Labov, William. 1981. Resolving the Neogrammarian controversy. *Language* 57: 267–308.
- Lambrecht, Knud. 1981. *Topic, antitopic and verb agreement in Nonstandard French*. Amsterdam: Benjamins.
- Langacker, Ronald. 1987. *Foundations of cognitive grammar. Vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- Lieven, Elena, Julian M. Pine and Gillian Baldwin. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24: 187–219.
- MacWhinney, Brian. 1978. *The acquisition of morphophonology*. Monographs of the Society for Research in Child Development, no. 174, vol. 43.
- Miller, George A. 1965. Introduction. In George K. Zipf, *The Psycho-Biology of Language*, i–x. Cambridge: The MIT Press.
- Moonwomon, Birch. 1992. The mechanism of lexical diffusion. Paper presented at the Annual Meeting of the Linguistic Society of America, January 1992, Philadelphia.
- Mowrey, Richard, and William Pagliuca. 1995. The reductive character of articulatory evolution. *Rivista di Linguistica* 7(1): 37–124.
- Ochs, Elinor, Emanuel A. Schegloff, and Sandra A. Thompson. 1996. *Introduction*. Ochs et al., eds., 1–51.
- Ochs, Elinor, Emanuel A. Schegloff, and Sandra A. Thompson, eds., 1996. *Interaction and Grammar*. Cambridge: Cambridge University Press. (Studies in Interactional Sociolinguistics 13)
- Ono, Tsuyoshi and Sandra A. Thompson. 1994. Unattached NPs in English conversation. *BLS* 20.
- Phillips, Betty S. 1998. Lexical diffusion is NOT lexical analogy. *Word* 49: 369–81.
- Pierrehumbert, Janet. 1994a. Knowledge of variation. *Papers from the Parasession on Variation*, 30th Meeting of the Chicago Linguistic Society, *CLS* 30: 232–56.
- Pierrehumbert, Janet. 1994b. Syllable structure and word structure: a study of triconsonantal clusters in English. In Patricia Keating, ed., *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, 168–90. Cambridge: Cambridge University Press.
- Renouf, A. and John McH. Sinclair. 1991. Collocational frameworks in English. In *English Corpus Linguistics*, K. Aijmer and B. Altenberg (eds.), 128–44. London: Longman.
- Rumelhart, D. and McClelland, J. 1986. On learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In *Parallel distributed processing: Explorations in the microstructure of cognition*, J. McClelland, D. Rumelhart, and the PDP Research Group (eds.). Cambridge, MA: MIT Press.
- Saussure, Ferdinand de. 1916. *Cours de linguistique général*. Paris: Payot.
- Scheibman, Joanne. 2000. *Structural Patterns of Subjectivity in American English Conversation*. Doctoral Dissertation. Albuquerque: Department of Linguistics, University of New Mexico.
- Schuchardt, Hugo. 1885. Über die Lautgesetze: Gegen die Junggrammatiker. [Translated in: *Schuchardt, the Neogrammarians and the transformational theory of phonological change*, Theo Vennemann and Terence Wilbur (eds.) 39–72. Frankfurt: Athenäum, 1972.]
- Sinclair, John McH. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

- Sinclair, John McH. 1992. Trust the text. In M. Davies and L. Ravelli (eds.) *Advances in Systemic Linguistics*, 5–19. London: Pinter.
- Stubbs, Michael. 1996. *Text and Corpus Analysis*. Oxford: Blackwell.
- Toolan, Michael. 1996. *Total Speech: An Integrational Linguistic Approach to Language*. Durham and London: Duke University Press.
- Traugott, Elizabeth. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65:31–55.
- Traugott, Elizabeth, and Ekkehard Koenig. 1991. The semantics-pragmatics of grammaticalization revisited. In E. Traugott and B. Heine (eds.) *Approaches to Grammaticalization, vol. 1*, 189–218. Amsterdam: John Benjamins.
- Vitevitch, Michael S., Paul A. Luce, Jan Charles-Luce, and David Kemmerer. 1997. Phonotactics and syllable stress: implications for the processing of spoken nonsense words. *Language and Speech* 40: 47–62.
- Wolf, George (ed.) 1999. *Special Issue on Integrational Linguistics in the Context of 20th Century Theories of Language*. *Language and Communication* 19,1. Oxford: Pergamon Press.
- Wray, Alison, and Michael R. Perkins. 2000. The functions of formulaic language: an integrated model. *Language and Communication* 20,1:2000, 1–28.
- Zipf, George K. 1965. *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Cambridge, Mass.: The MIT Press. (First published in 1935 by Houghton Mifflin.)