

GRAMMAR IS GRAMMAR AND USAGE IS USAGE

FREDERICK J. NEWMAYER

University of Washington

A number of disparate approaches to language, ranging from cognitive linguistics to stochastic implementations of optimality theory, have challenged the classical distinction between knowledge of language and use of language. Supporters of such approaches point to the functional motivation of grammatical structure, language users' sensitivity to the frequency of occurrence of grammatical elements, and the great disparity between sentences that grammars generate and speakers' actual utterances. In this article I defend the classical position, and provide evidence from a number of sources that speakers mentally represent full grammatical structure, however fragmentary their utterances might be. The article also questions the relevance of most corpus-based frequency and probability studies to models of individual grammatical competence. I propose a scenario for the origins and evolution of language that helps to explain why grammar and usage are as distinct as they are.*

1. INTRODUCTION. My goal in this article is to argue in favor of the classical Saussurean position with respect to the relationship between knowledge of language and use of language. That is, the article provides evidence in support of the idea that the mental grammar contributes to language use, but that usage, frequency, and so on are not represented in the grammar itself. To do so, I necessarily take a wide-ranging approach, describing current anti-Saussurean *USAGE-BASED MODELS* and attempting to account for their popularity, and then defending the classical position with arguments based on the compatibility of formal grammar and functional explanation, the failure of connectionism to provide an alternative to formal grammar, and the fact that speakers mentally represent full argument structure. Further, I argue that the mental grammar is only one of many systems that drive usage, since grammars are not actually well designed to meet language users' 'needs'. I explore as well the relevance for the main theme of the article the fact that grammatical change is syntagmatic, rather than paradigmatic. An argument against stochastic grammars is presented and one against the idea that grammars are fragile temporary objects. Finally, I propose an evolutionary scenario that makes sense of the grammar-use distinction.

2. USAGE-BASED MODELS OF GRAMMAR. The late 1960s witnessed the birth of the approach to grammar called *GENERATIVE SEMANTICS*. This approach promised to revolutionize the field of transformational generative grammar, itself barely a decade old at the time. With each passing year, generative semanticists declared that some seemingly well-established boundary was nothing but an illusion. The framework began in 1967, when George Lakoff and John R. Ross challenged the existence of the level of Deep Structure, and with it the boundary between syntax and semantics (Lakoff & Ross 1967/1976). The following year McCawley (1968) argued that syntactic and lexical structures were formal objects of the same sort. Then in quick succession, the dividing line between semantics and pragmatics (Lakoff 1970/1972), grammaticality and ungrammaticality (Lakoff 1973), category membership and nonmembership (Ross 1973a,b), and, finally, grammar and usage (Lakoff 1974) were all cast into doubt. At the same time, many sociolinguists were proposing models in which statistical facts about the speech community were incorporated into grammatical rules (Labov 1969,

* This article is a slightly revised version of my Linguistic Society of America presidential address, presented in Atlanta in January 2003. I would like to thank Brian Joseph, Richard Oehrle, Thomas Wasow, and an anonymous referee for their comments on an earlier version of the entire manuscript.

1972). By the late 1970s, however, generative semantics, for reasons I have discussed elsewhere (Newmeyer 1986), had all but disappeared. Most syntacticians had re-embraced the boundaries whose demise had been heralded only a few years earlier.

The last decade has seen the resurgence of many of the same ideas that were the hallmark of generative semantics. In particular, most of the ways of looking at form and meaning that fall under the rubric of COGNITIVE LINGUISTICS have reasserted—albeit in different form—the bulk of the ideas that characterized generative semantics. Langacker (1987:494) coined the term USAGE-BASED MODEL to refer to those approaches that reject a sharp distinction between language knowledge and language use. My impression is that more linguists around the world do cognitive linguistics than do generative grammar. Many functional linguists share the view of a usage-based model; indeed, the dividing line between cognitive linguistics and functional linguistics has never been sharp. The following quote from two prominent functionalists gives the flavor of what is implicit in a usage-based model:

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; 1988; 1998), understood as an ongoing process of STRUCTURATION (Giddens 1984) . . . [E]mergent structures are unstable and manifested stochastically . . . 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. (Bybee & Hopper 2001a: 2–3)

Some of these ideas have been adopted by syntacticians working in the generative tradition. In particular, language-particular and crosslinguistic statistical regularities have been incorporated into stochastic implementations of optimality theory:

The same categorical phenomena which are attributed to hard grammatical constraints in some languages continue to show up as statistical preferences in other languages, motivating a grammatical model that can account for soft constraints . . . [T]he stochastic OT framework can provide an explicit and unifying framework for these phenomena in syntax. (Bresnan et al. 2001:15; see also the discussion in Wasow 2002)

I believe that the great majority of psycholinguists around the world consider the competence-performance dichotomy to be fundamentally wrongheaded. Usage-based approaches have swept natural language processing as well. I am quite sure that Christopher Manning is right when he writes that ‘[d]uring the last 15 years, there has been a sea change in natural language processing (NLP), with the majority of the field turning to the use of machine learning methods, particularly probabilistic models learned from richly annotated training data, rather than relying on hand-crafted grammar models’ (Manning 2002b:441).

3. THE APPEAL OF USAGE-BASED MODELS. The obvious question to ask is why there has been a change of mind among many theoretical linguists. If the ideas that characterized generative semantics were laid to rest in 1970s, then why are they back again? There are several reasons. First and most importantly, there is the evidence that has mounted in the past quarter-century that significant aspects of grammars are motivated by considerations of use. Functional linguists and generative linguists with a functional bent have provided (to my mind) incontrovertible evidence that grammars are shaped in part by performance considerations. While there is no space to review this rich literature, the following publications stand out in my mind as among the most noteworthy: Bybee 1985, Comrie 1989, Croft 1990, Givón 1995, Haspelmath 1999, Hawkins

1994, Hopper & Thompson 1980.¹ I find that the most compelling examples of function affecting form pertain to the shaping of grammars by on-line processing pressure. So take the tendency of heads consistently to precede complements or to follow complements. One might be tempted to simply declare a head-parameter provided by Universal Grammar and leave it at that. There very well MIGHT BE a head parameter, as far as descriptively adequate grammars are concerned. But there is a lot more to the story. Consider a VO language like English, where heads typically precede complements:

- (1) V-NP, P-NP, A-of-NP, N-of-NP

In each case a 'lighter' head precedes a 'heavier' complement. But the light-before-heavy tendency in the grammar involves far more than the head-complement relation. For example, the canonical order of VP constituents is relentlessly lighter-to-heavier, as in 2.

- (2) _{VP}[V - NP - PP - CP] (*convince my students of the fact that all grammars leak*)

Also notice that single adjectives and participles can appear in pre-head position.

- (3) a. a silly proposal
b. the ticking clock

But if these adjectives and participles themselves have complements, the complements have to appear in post-head position.

- (4) a. *a sillier than any I've ever seen proposal
b. a proposal sillier than any I've ever seen
(5) a. *the ticking away the hours clock
b. the clock ticking away the hours

Many more examples of this phenomenon could be provided. As Hawkins (1994) has shown, all of these facts follow from performance-based pressure to shorten the recognition time for phrasal constituents.² Evidence for its performance basis is simple. Where speakers have a choice in a VO-type language, they tend to put shorter before longer constituents. So, except for cases in which there is a strong lexical relation between V and P, PP's can typically occur in any order after the verb.

- (6) a. Mary talked to John about Sue.
b. Mary talked to Sue about John.

But all other things being equal, the greater the length differential between the two PP's, the more likely speakers will be to put the shorter one first.³ Interestingly, Hawkins's approach makes precisely the opposite length and ordering predictions for head-final languages. And to be sure, there is a heavy-before-light effect in those languages, both in language use and in the grammar itself. So there is no question in my mind that grammars have been shaped by processing considerations—that is, by language in use.

¹ For detailed discussion of the mechanisms by which performance considerations can lead to changes in grammatical competence and of the scale at which the interaction takes place, see Newmeyer 1998, 2001.

² An anonymous referee points out that English allows 'heavy' sequences of adjectives in prenominal position: *the big, black, hairy, ferocious dog*. In fact, Hawkins's approach predicts that the short-before-long effect will be much less pronounced if the elements that contribute to prehead heaviness are all of the same grammatical category. Seeming prehead phrases such as *easy to please in an easy to please client* are actually complex lexical items (Nanni 1978, Roeper & Siegel 1978).

³ The discourse status of the elements involved also plays a role in ordering (see Arnold et al. 2000, Hawkins 2003).

Related to work showing that grammars are externally motivated is the increasing realization that language users are sensitive to the FREQUENCY of grammatical forms (see especially the papers in Bybee & Hopper 2001b). This notion has in fact been suggested for well over a century in the domain of phonological change, dating from at least Schuchardt (1885/1972), though the exact way in which word frequency interacts with sound change is controversial.⁴ In a somewhat parallel fashion, though not necessarily the reflection of a sound change per se, my own pronunciation of words with orthographic ⟨o⟩ between or next to consonants varies in a way connected to a word's frequency, even when the same consonants are involved. Table 1 illustrates.

| /ɔw/ | /a/ |
|---------------------|-------------------------------|
| on, off | honor, offal, don, doff, Goth |
| dog | frog, log, bog |
| loss, boss | floss, dross |
| strong, song, wrong | gong, tong |
| cost, frost | Pentecost |

TABLE 1. Some vowel contrasts in the author's English.

Frequency effects have been pointed to at levels higher than the phonological. It is well accepted, I think, that frequently occurring forms tend to be shorter than less frequently occurring ones (Zipf 1935, 1949, Horn 1984, 1993). And a crucial component of David Lightfoot's analysis of the rise of VO word order in Middle English is language-learner sensitivity to declining token frequency of OV order (Lightfoot 1991). Again, such facts might be taken to suggest that grammar and use are too intertwined to be separated one from the other.

Reinforcing skepticism about classical generative models is the disparity between sentences generated by these grammars and actual utterances produced by language users. This disparity has led some linguists to conclude that grammar itself bears no relation to the proposition-like structures posited by formal linguists; structures specified by formal rules that take the sentence to be the basic unit of grammar, where sentences are in a rough mapping with propositions, verbs with predicates, and noun phrases with logical arguments. The priority of the sentence is dismissed by some critics of the generative program as a carryover from the Western logical tradition, reinforced by the conventions of written language (see especially Harris 1980, 1981).

Why would anyone draw that conclusion? Well, the argument goes that in ACTUAL SPEECH speakers rarely utter sentences with a subject, a verb, and an object, where the two arguments are full lexical items, even though that is what grammars generate. Rather, what one finds the most is what Du Bois (1987) calls 'preferred argument structure'. Most utterances consist of a verb with one full argument, which is either the subject of an intransitive verb or the object of a transitive verb. Other arguments are either reduced to clitic or affix status or omitted entirely. Examples are provided in 7.

⁴ Studies of lexical diffusion are relevant here (see Wang 1969, Chen & Wang 1975, Wang 1977 for discussion and references), though there is controversy as to whether frequent items are affected first or last, as discussed by Phillips (1984). There are some linguists who even doubt the existence of lexical diffusion as it pertains to sound change, though they recognize a diffusionary effect in the spread of pronunciation changes (e.g. Joseph 2001).

(7) Examples of PREFERRED ARGUMENT STRUCTURE

| | |
|--|--|
| Cayuga (Iroquoian) | 1–2% of clauses contain 3 major constituents (Mithun 1987) |
| Chamorro (Austronesian) | 10% of transitives have 2 lexical arguments (Scancarelli 1985) |
| Hebrew (Semitic) | 93% of transitive clauses lack a subject NP (Smith 1996) |
| French (Romance) | French preferred clause structure is [(COMP) clitic + Verb (X)]; only 3% of clauses contain lexical subjects (Lambrecht 1987) |
| German (Germanic) | even ditransitive verbs in spoken discourse tend to follow preferred argument structure (Schuetze-Coburn 1987) |
| Huallaga Quechua (Andean) | in one corpus, only 8% of sentences contained both a noun subject and a noun object (Weber 1989) |
| Coos (Penutian) | 2–3% of clauses contain 3 major constituents (Mithun 1987) |
| Mam (Mayan) | 1% of clauses have 2 lexical arguments (England 1988) |
| Malay (Austronesian) | ‘Malay is thus similar to what Du Bois (1985) has described for Sacapultec Maya: it has a ‘‘Preferred Argument Structure’’ ’ (Hopper 1988:126) |
| Ngandi (Australian) | 2% of clauses contain 3 major constituents (Mithun 1987) |
| ‘O’odham = Papago (Uto- Aztecan) | only 9% of transitives have 2 overt arguments (Payne 1992) |
| Rama (Chibchan) | transitive clauses with 2 NPs are rare (Craig 1987) |
| Sacapultec (Mayan) | in connected discourse, only 1.1% of clauses have 2 lexical arguments (Du Bois 1985, 1987) |
| Yagua (Peba- Yaguan) | in a corpus of 1516 clauses, only 3% contained both a noun subject and a noun object (Payne 1990) |

Even English, which is non-null subject and considered rigidly SVO, manifests preferred argument structure. A corpus of 20,794 sentences (from telephone conversations) included only 5,975 (29%) that were SVO (Dick & Elman 2001; see also Thompson & Hopper 2001 for similar figures). So if real speech, the argument goes, is not propositional, then grammars should not be either. Many more studies of language in use (e.g. Fox 1994, Thompson & Hopper 2001) lead to similar conclusions.

Finally, I think that the appeal of usage-based models is based in part on the rise in the past fifteen years of an approach to the human mind that seems to allow no place for the algebraic autonomous grammars of classical generative grammar. I am referring of course to connectionism (PDP models). If all that we have are stored activation weights, which themselves are no more than predispositions to behave, and connections among them, then usage-based approaches to grammar would seem to follow as a matter of course.

4. FUNCTIONAL EXPLANATION IS FULLY COMPATIBLE WITH FORMAL GENERATIVE GRAMMAR. I think that some arguments in favor of usage-based models can be dismissed right away. Most importantly, one that can be laid to rest is based on the fact that since properties of grammars are functionally motivated, grammar and use are necessarily

inextricable. Such a view seems to assume that once a system is characterized as discrete and algebraic, a functional explanation of that system (or its properties) becomes impossible. But that is simply not true. Indeed, it seems to be only linguists who have this curious idea. In every other domain that I am aware of, formal and functional accounts are taken as complementary, rather than contradictory.

The point can be illustrated with a look at a couple of other systems (which are discussed in more detail in Newmeyer 1998:ch. 3). Consider the game of chess. One could not ask for a better formal system. There exists a finite number of discrete rules; given the layout of board, the pieces, and their moves, every possible game of chess can be ‘generated’. But functional considerations went into the design of system. Presumably it was designed in such a way as to make it a satisfying pastime. And external factors can change the system. However unlikely, a decree from the International Chess Authority could change the rules for castling. Furthermore, in any game of chess, the moves are subject to conscious will of the players, just as in the act of speaking, the conscious decision of the speaker plays a central role. So chess is a formal system and explained functionally.

A more biological analogy can be provided by any bodily organ. The liver, say, can be described as an autonomous structural system. But still it has been shaped by its function and use. This organ evolved in response to selective pressure for a more efficient role in digestion. And it can be affected by external factors. A lifetime of heavy drinking can alter its structure. So I simply do not see any merit in pointing to functional explanation as a wedge against the classic view of what grammars are like.

By the way, I regard the assumption that much of grammatical structure is motivated by external functional pressure as being a fairly uncontroversial one, even among the most doctrinaire formal linguists. Certainly Chomsky has never questioned it. As long ago as 1975 he wrote

Surely there are significant connections between structure and function; this is not and has never been in doubt . . . Searle argues that ‘it is reasonable to suppose that the needs of communication influenced [language] structure’. I agree. (Chomsky 1975:56–58)

More recently, in *The minimalist program* (Chomsky 1995) and in subsequent work (Chomsky 2000, 2001), he suggests that displacement phenomena—that is, movement rules—probably exist to facilitate language use, both in terms of parsing needs and the demands of information structure. So the issue is not whether grammars have functional motivation, but where and how much, and the centrality of focusing on this motivation in one’s research program.⁵

5. CONNECTIONISM AND USAGE-BASED MODELS OF GRAMMAR. As far as connectionism is concerned, there is not much to say. We have come a long way from the days when connectionism was little more than behaviorism on the computer. To be sure, the very earliest connectionist models from the 1980s and early 1990s (e.g. Rumelhart & McClelland 1986) were hopeless at capturing even the most basic aspects of grammar, such as long-distance dependencies, category-sensitive processes, structure-dependence, and so on. The problem is that now anything can be called ‘connectionist’, as long as it involves modeling on a network. In particular, current connectionist models have no trouble implementing rules, and they can even, by virtue of prespecified connection weights, mimic a rich innate component to grammar. I take that point as being uncontroversial. Consider what some connectionists themselves say (or those who are in general sympathetic to the endeavor).

⁵ For Chomsky, of course, such a focus has no centrality.

Connectionist models do indeed implement rules. (Elman et al. 1997:176)

Thus a commitment to the PDP Principles . . . does not per se constitute a commitment regarding the degree to which discreteness, modularity, or innate learning applies to human cognition. (Smolensky 1999:592)

The emphasis in the connectionist sentence-processing literature on distributed representation and emergence of grammar from such systems can easily obscure the often close relations between connectionist and symbolic systems . . . Connectionism is no more intrinsically non-modular than any other approach, and many connectionists . . . have explicitly endorsed modular architectures of various kinds. (Steedman 1999:615)

And despite that, connectionist modeling still has little to show for itself when it comes to handling grammatical facts. Again, I let sympathizers speak for themselves.

The [connectionist] approach [to language] is new and there are as yet few solid results in hand. (Seidenberg 1997:1602)

And this was after a decade of PDP work on language!

I realize how enormous the gap is between existing PDP models and a system that would approximate the actual complexity of linguistic structure, even in limited domains. (Langacker 2000:6)

Despite these grand pretensions, the reality of connectionist modeling is more sober and modest. In fact, much of the work to date has focused on the learning of narrow aspects of inflectional morphology in languages like English and German. (MacWhinney 2000:125)

So for various reasons, I do not think that one needs to dwell on the achievements of connectionism as somehow tipping the scales in favor of usage-based models.

6. FULL GRAMMATICAL STRUCTURE IS COGNITIVELY REPRESENTED. Let us now turn to what I take to be greater challenges to the classical Saussurean position. Consider the disparity between what are characterized as grammatical sentences and the sorts of utterances that speakers actually make. Does the fact that most utterances reflect preferred argument structure rather than full argument structure militate against grammars in which predicate-argument structure and so on play a central role? No, it does not.

I begin with some typological evidence. Since Greenberg 1963 it has been customary to divide languages into their basic ordering of subject, verb, and object. From this ordering, certain predictions tend to follow. SOV languages, for example, are more likely to have postpositions than prepositions. SVO languages are more likely to have WH-movement than SOV languages, and VSO languages are even more likely to have it than SVO languages. But as we have seen, very few utterances in natural language actually have the full subject, verb, and object. If speakers do not mentally represent the full propositional structure, then a prediction follows. The prediction is that the ordering of full subject, verb, and object should be irrelevant to typology. But that prediction is false, as a look at French indicates. French is SVO when the object is lexical, but SOV when the object is pronominal.

- (8) a. Marie voit Jean.
b. Marie le voit.

Text counts show that sentences like 8b are vastly more likely to be uttered than those like 8a. But French is archotypically a VO language in its typological behavior. In other words, in this case actual language use is irrelevant. What is important is the ordering of elements in sentences with full propositional and lexical structure, rare as they are in actual discourse. The reason is that it is this full structure that is called upon in speech production. According to Levelt (1989), a central part of planning a speech act involves retrieving lexical information, what he calls 'lemmas', essentially predi-

cates and their argument structure. In other words, for transitive verbs like *hit*, *know*, *eat*, and so on the speaker has a mental representation of the full argument structure of the sentence: ‘Lemma structure plays a central role in the generation of surface structure. In particular, the main verb dictates what arguments have to be checked in the message, and which grammatical functions will be assigned to them’ (Levelt 1989: 244).

The FORMULATOR, the formulating component of speech production, takes this information as input. Because of that, sentences with full argument structure are psychologically more basic than others, even though efficient discourse packaging keeps them from being used very often. But for typological purposes, it does not matter that fully elaborated sentences are rarely actually used. It is the most frequently used canonical ordering of subject, verb, and object that drives typology, not the most frequently used utterance type in general. The speaker of French might rarely utter sentences like 8a, but he or she does so far more often than those like 9.

(9) *Marie Jean voit.

And because of that, French behaves typologically like a VO language, not like an OV language.⁶

The way that sentence fragments are processed also points to the centrality of fully specified grammatical representations, as work such as Pope 1971 and Morgan 1973 reveals. Consider some possible answers to the question in 10. Those in 11a–c are possible, but not those in 11d–e.

(10) Who does John_i want to shave?

- (11) a. Himself_i
 b. Him_j
 c. Me
 d. *Myself
 e. *Him_i

The generalization, of course, is that the possible pronoun corresponds to the one usable in full sentences, with all arguments expressed (12).

- (12) a. John_i wants to shave himself_i.
 b. John_i wants to shave him_j.
 c. John_i wants to shave me.
 d. *John_i wants to shave myself.
 e. *John_i wants to shave him_i.

In other words, whatever one might do in actual speech, one’s cognitive representation embodies all the arguments and the principles for assigning the proper pronominal form to the direct object.⁷ Here are a couple of other examples illustrating the same point.

⁶ Dryer (1997) proposes replacing Greenberg’s word order typology, which appeals to the basic ordering of S, V, and O for each language, with two two-way parameters, SV vs. VS and OV vs. VO. Since sentences with just V and S or just V and O are common in discourse, Dryer points out that his system does not rely on discourse-rare sentence types with all three arguments expressed. For a defense of the Greenbergian system, see Newmeyer 2003b.

⁷ Richard Oehrle and Thomas Wasow have pointed out to me that fragment types exist whose grammatical properties do not follow directly from their corresponding full sentences. For example, Oehrle notes that *Clinton* can be understood as the antecedent of *himself* in (ib), even though the coreference depicted in (ii) is impossible.

- (i) a. Who did Clinton say that Cuomo planned to nominate?
 b. Himself.
 (ii) *Clinton_i said that Cuomo planned to nominate himself_i.

The verb *think* in English takes a finite complement, but not a gerundive or infinitival one.

- (13) a. Mary thinks that Sue has finally solved the problem.
 b. *Mary thinks Sue's having finally solved the problem.
 c. *Mary thinks Sue to have finally solved the problem.

Notice the possible answers to 14: only 15a is possible, not 15b or 15c.

- (14) What does Mary think?
 (15) a. Sue has finally solved the problem.
 b. *Sue's having finally solved the problem.
 c. *Sue to have finally solved the problem.

The internal structure of NP arguments has to be preserved as well, as 16a–c show.

- (16) a. Does Alice like the soprano?
 b. No, the tenor.
 c. *No, tenor.

These observations all lead to the conclusion that speakers mentally represent the full grammatical structure, even if they utter only fragments.

It is worth pointing out that the disparity between knowledge of grammar and actual usage is taken for granted by most researchers of infant and child speech. Clearly, infants know more about grammar than is reflected by their utterances. This has been demonstrated by series of experiments carried out by Hirsh-Pasek and Golinkoff. They report (Hirsh-Pasek & Golinkoff 1996) that

- (a) one-word speakers between thirteen and fifteen months know that words presented in strings are not isolated units, but are part of larger constituents;
 (b) One-word speakers between sixteen and nineteen months recognize the significance of word order in the sentences that they hear;
 (c) Twenty-eight month-old children who have productive vocabularies of approximately 315 words and who are speaking in four-word sentences can use a verb's argument structure to predict verb meaning.

The grammatical principles that very young children have assimilated can be extremely complex and abstract. Consider the acquisition of anaphoric binding. Crain and McKee (1986) have shown that even children as young as two years old understand the coreference possibilities in 17a–d.

- (17) a. The Ninja Turtle danced while he ate pizza.
 b. While he danced, the Ninja Turtle ate pizza.
 c. His archrival danced while the Ninja Turtle ate pizza.
 d. He danced while the Ninja Turtle ate pizza.

They recognize that in 17a–b *he* can be coreferential with *the Ninja Turtle*, in 17c *his* can be coreferential with *the Ninja Turtle*, but that in 17d *he* and *the Ninja Turtle* cannot corefer.

Examples like 17 and like Hirsh-Pasek and Golinkoff's findings are often given to support the idea that much of grammar is innate. They very well MIGHT support that idea, but that is not my purpose for giving them here. This article has nothing to contribute to the issue of innateness. The importance of such examples is that they

Such facts show that there is more involved in interpretation of fragments than can be predicted from their presumptive underlying sources. However, the facts of 10–12 are sufficient to make the point that language users do mentally represent these sources.

support a much weaker claim than the innateness of grammatical structure, though still a controversial one. They indicate quite clearly that the understanding of knowledge of grammar involves going beyond an examination of language in use.

Let us consider another piece of evidence that language users have representations that are not predictable from usage-based facts about language. Generative grammarians have long been castigated by other linguists for working with sentences that they make up out of their heads, rather than those taken from actual texts. Now, there are many pros and many cons to the use of introspective data and it is not my purpose to review them here (for a good discussion, see Schütze 1996). However it is worth pointing out a remarkable fact about the human language faculty that would never have been unearthed if one just confined one's attention to usage. Speakers have the remarkable ability to make reliable judgments about sentence types that they only rarely hear or utter. Take sentences with parasitic gaps, as in 18.

(18) This is the paper_i that I filed_i before reading_i.

I believe that these are rare in actual speech, though I do not know of any statistical studies to confirm that claim. But I doubt that there exists an adult speaker of English who does not know that 18 is a better sentence than 19a–b, despite their superficial similarities.

(19) a. *I filed the paper_i before reading_i.

b. *This is the paper_i that I filed the notes before reading_i.

'Useless' as it is to know the facts surrounding 18 and 19, we know them anyway. Recent experimental work has confirmed that speakers can make reliable introspective judgments, even about rarely occurring sentence types. Cowart (1997) took some constructions that have loomed large in theoretical discussions (the examples in 20 to 22) and showed that there was a stable pattern of response to them among his subjects.

(20) Subjacency

- a. Why did the Duchess sell a portrait of Max?
- b. Who did the Duchess sell a portrait of?
- c. Who did the Duchess sell the portrait of?
- d. Who did the Duchess sell Max's portrait of?

(21) That-trace phenomena

- a. I wonder who you think likes John.
- b. I wonder who you think John likes.
- c. I wonder who you think that likes John.
- d. I wonder who you think that John likes.

(22) Coordination and binding theory

- a. Cathy's parents require that Paul support himself.
- b. Paul requires that Cathy's parents support himself.
- c. Cathy's parents require that Paul support himself and the child.
- d. Paul requires that Cathy's parents support himself and the child.

These are not particularly common sentence types in use, and yet experimental subjects are quite consistent in how they judge them.⁸

Along the same lines, McDaniel and Cowart (1999) found that subjects can reliably rate sentences like those in 23 in terms of degree of acceptability.

(23) Resumptive pronouns

- a. That is the girl that I wonder when met you.

⁸ The results reported by Cowart do not always support the grammaticality assignments found in the generative literature, but that is another story.

- b. That is the girl that I wonder when she met you.
- c. That is the girl that I wonder when you met.
- d. That is the girl that I wonder when you met her.

These results indicate that there is a lot more to grammar than can be predicted from use in naturally occurring discourse. More importantly, they suggest that the human language faculty is designed—at least in part—for something other than communication, a point to which I return below.

7. THE MENTAL GRAMMAR: ONE SYSTEM THAT DRIVES USAGE. No generative grammarian ever claimed that sentences generated by the grammar should be expected to reveal directly what language users are likely to say. This must be true for an obvious reason, namely, that knowledge of grammatical structure is only one of many systems that underlie usage. So it will not do as a refutation of formal grammar to find some generalization that the grammar does not encompass. The explanation of that generalization might lie somewhere else. Lakoff and Johnson's mistake in their book *Philosophy in the flesh* (Lakoff & Johnson 1999) was to assume that any generalization about usage is necessarily a matter for grammar to handle. They pointed out that deictic locatives and rhetorical questions like 24a–b have traditionally been assumed to occur only as main clauses.

- (24) a. Here comes the bus! (deictic locative)
- b. Who on earth can stop Jordan? (rhetorical question)

They show, however, that such sentences occur in subordinate clauses introduced by *because* (25), though not in those introduced by *if* (26).

- (25) a. I'm leaving because here comes my bus.
- b. The Bulls are going to win because who on earth can stop Jordan?
- (26) a. *I'm leaving if here comes my bus.
- b. *The Bulls are going to win if who on earth can stop Jordan?

Lakoff and Johnson's reasonable generalization is that deictic locatives and rhetorical questions can occur in certain subordinate clauses because they convey statement speech acts. Since *because*-clauses express a reason and the statement speech act is the reason, deictic locatives and rhetorical questions can occur after *because*. Lakoff and Johnson conclude that since a pragmatic generalization unites the syntactic constructions, it is wrong-headed to dissociate grammar from usage. But their solution is too limited and construction-specific. What we have here is a generalization about what constitutes a coherent discourse. Notice that 27a is as good as 25a (and for the same reason), while 27b is as bad as 26a (and for the same reason).

- (27) a. I'm leaving. Do you know why? Here comes my bus.
- b. #I'm leaving if the following condition is met. Here comes my bus.

In other words, we are not dealing with a grammatical generalization at all. English grammar does not distinguish between *because*- and *if*-clauses with respect to the issues that concern us. A condition (presumably universal) on coherent discourse is at work here.⁹

⁹ Given the fact that there are more grammatical restrictions obtaining in subordinate clauses than in main clauses (Ross 1973c, Emonds 1976), some well-formed sequences of simple sentences will correspond to ungrammatical sentences with subordinate clauses. An example is the following, pointed out to me by Richard Oehrle.

- (i) *Because here comes my bus, I'm leaving.
- (ii) Here comes my bus. I'm leaving.

When one thinks about the nature of speech, it is pretty obvious why other systems besides grammar per se have to take over a big part of the job of communication. The transmission rate of human speech is painfully slow. According to Bill Poser (cited as a personal communication in Levinson 2000:382), it is less than 100 bits per second—compared to the thousands that the personal computer on one’s desk can manage. A consequence is that speakers have to pack as much as they can into as short a time as they can, leading to most utterances being full of grammatical ambiguity. For example, sentence 28 was calculated by Martin and colleagues (1987) to have 455 parses.

- (28) List the sales of products produced in 1973 with the products produced in 1972.

For that reason, humans have developed complex systems of inference and implicature, conveyed meanings, and so on. Grammar is such a poor reflection of usage because we have many more meanings to convey than could ever be supported by our grammatical resources in a reasonable period of time. Stephen Levinson phrased it beautifully: ‘[I]nference is cheap, articulation expensive, and thus the design requirements are for a system that maximizes inference’ (Levinson 2000:29).

8. GRAMMARS NOT TAILOR-MADE TO SERVE USERS’ ‘NEEDS’. The role played by extragrammatical systems in communication means that grammar per se is not always going to tailor itself to our communicative needs. But a standard assumption in a lot of functionalist writing is that grammar DOES oblige itself in that way. This assumption has been enshrined in a famous dictum from Jack Du Bois.

Grammars provide the most economical coding mechanism . . . for those speech functions which speakers most often NEED to perform. More succinctly: Grammars code best what speakers do most. (Du Bois 1985:362–63; emphasis added)

That is much too strong a position. The problem is that appeals to ‘need’ have a post-facto feel to them. One observes a generalization and comes up with a plausible story to account for it. But there is nothing PREDICTIVE about accounts that say that grammars encode whatever because that is what speakers need grammars to do. One can imagine any number of things that it would be useful for grammars to do, but which they never do. We need a theory, which is now lacking, of why some seemingly useful features result in grammatical coding, and some do not.

It is worth providing a few examples of how we rely on extragrammatical systems to take over the job of conveying meaning, thereby exempting the grammar from the burden of meeting our ‘needs’. Consider first the inclusive/exclusive pronoun distinction.

- (29) Inclusive-exclusive pronoun distinction in Washo (Jacobsen 1980)

| | SG. DUAL | PL. |
|---------------|-------------------------------|----------------------------|
| 1st exclusive | lé’ léši (= I and one other) | léw (= I and others) |
| 1st inclusive | léšiši (= I and you [sg.]) | léwhu (= I and you [pl.]) |

According to Nichols (1992), only a minority of languages (about 42%) make this distinction. And it is heavily areal—only in Australia, Oceania, and South America do more than 60% of languages manifest it. Yet the distinction is quite useful from an ambiguity-reduction perspective. We all have been in situations where we or someone else has said ‘We are going to do X, Y, and Z’ and it has not been clear whether the person addressed was included in the ‘we’ or not. So a distinction that we ‘need’ is not generally lexicalized.

One could make the same point about the falling together of second person singular and plural pronouns in English (where it is total) and in other European languages

(where it affects the polite form). How useful is that? Consider also the fact that a majority of the world's languages are null subject (Gilligan 1987). That might make language faster to use, even though it creates more potential ambiguity in doing so. Or take deictic systems. They do not seem particularly designed to meet language users' needs either. They are typically organized in terms of distance from speaker or hearer. The height of an object in relation to the speech participants seems like a useful distinction for grammars to make. But according to Anderson and Keenan (1985), only a small handful of languages encode this distinction grammatically. Speakers could also avoid potential ambiguity by means of a grammatical marker specifying whether an object is on their left side or their right side. But Hawkins (1988) writes that no language grammaticalizes this distinction. Another striking fact, reported in Talmy 1985/2000, is that no language has grammatical means (inflectional morphology, incorporations, etc.) for encoding an event structure that is independent of that encoded by the lexical categories in the sentence. In other words, no matter how useful it would be to provide grammars with the possibility of sentences like 30 with meanings like 31 and 32, no language allows that possibility.

- (30) The chair broke-ka.
 (31) The chair broke and I'm currently bored.
 (32) The chair broke and it was raining yesterday.

9. GRAMMATICAL CHANGE IS SYNTAGMATIC, NOT PARADIGMATIC. It is a good question whether the drive to reduce ambiguity, as useful as it might be to languages users—and therefore language use—is ever much of a driving force in grammatical change. Labov (1994) has explicitly argued that it is not. He based his conclusion on the observation that in actual speech one variant is rarely chosen over another for paradigmatic reasons, that is, in a fashion designed to preserve information. The choice is mechanical and syntagmatic—for example, phonetic conditioning and repetition of the preceding structure. Bill Croft has argued along similar lines in his book *Explaining language change*. He also stresses syntagmatic over paradigmatic change:

Form-function reanalysis [one of Croft's principal mechanisms of change—*FJN*] is syntagmatic: it arises from the (re)mapping of form-function relations of combinations of syntactic units and semantic components. The process MAY NEVERTHELESS HAVE AN APPARENTLY PARADIGMATIC RESULT, for example, a change of meaning of a syntactic unit. (Croft 2000:120; emphasis added)

In support of Labov and Croft, I cite what looks like a compelling example of paradigmatic pressure on morphosyntax and show how the facts can be reanalyzed in syntagmatic terms. Faltz (1977/1985) and Comrie (1998) point out that if a language has first and second person reflexives, it also has third person reflexives, as 33 illustrates.

(33) Occurrence of distinctive reflexives

| | 3RD PERSON | 1ST/2ND PERSON |
|-------------|------------|----------------|
| English | yes | yes |
| Old English | no | no |
| French | yes | no |
| * | no | yes |

Faltz's and Comrie's explanation for 33 is based on the idea that first and second person referents are unique. But third person referents are open-ended. In principle, a third person referent could be any entity other than the speaker or the hearer. So it would seem to be more useful to have third person reflexives, since they narrow down the

class of possible referents. Hence it appears that grammars are serving our needs by reducing potential ambiguity.

I can offer a syntagmatic explanation of these facts, that is, an explanation that does not involve problematic appeals to ambiguity-reduction. In languages that have reflexive pronouns in all three persons, third person reflexives are used more frequently than first and second. Consider English. In a million-word collection of British English texts, third person singular reflexives were 5.8 times more likely to occur than first person and 10.5 times more likely to occur than second person (Table 2 gives the facts).

| REFLEXIVE PRONOUN | OCCURRENCES IN CORPUS |
|---------------------|-----------------------|
| myself | 169 |
| yourself | 94 |
| himself | 511 |
| herself | 203 |
| itself | 272 |
| TOTAL 3RD PERS. SG. | 986 |

TABLE 2. Reflexive pronoun occurrence in English (Johansson & Hofland 1989).

Language users (for whatever reason) more frequently use identical subjects and objects in the third person than in the first or second. Given that more frequently appealed to concepts are more likely to be lexicalized than those that are less frequently appealed to, the implicational relationship among reflexive pronouns follows automatically. There is no need to appeal to ambiguity-reducing ‘usefulness’. Also, it is worth asking how much ambiguity is reduced by a third person reflexive anyway. It eliminates one possible referent for the object, leaving an indefinite number of possibilities remaining.

In summary, we have grammar and we have usage. Grammar supports usage, but there is a world of difference between what a grammar is and what we do—and need to do—when we speak.

10. SOME ARGUMENTS AGAINST STOCHASTIC GRAMMARS. Because of the divergence between grammar and usage, one needs to be very careful about the use made of corpora in grammatical analysis, and particularly the conclusions derived from the statistical information that these corpora reveal.¹⁰ Now, for some purposes, as we have seen, statistical information can be extremely valuable. Corpora reveal broad typological features of language that any theory of language variation, use, and change has to address. Two examples are the prevalence of preferred argument structure and the predominance of third person reflexives. And it goes without saying that facts drawn from corpora are essential for engineering applications of linguistics. But it is a long way from there to the conclusion that corpus-derived statistical information is relevant to the nature of the grammar of any individual speaker, and in particular to the conclusion that grammars should be constructed with probabilities tied to constructions, constraints, rules, or whatever.

¹⁰ I am indebted to Brady Clark, Christopher Manning, and Gabriel Webster for their helpful discussions with me on the topic of this section and especially for their good humor in being willing to take part in an exchange with a colleague like me whose ideas are fundamentally opposed to their own (at least as far as the first two individuals cited are concerned). The literature defending stochastic approaches to linguistic phenomena is vast—Jelinek 1997, Bod 1998, Manning & Schütze 1999, Jurafsky & Martin 2000, and Bod et al. 2003 are important recent book-length works. A comprehensive refutation of this literature is yet to be written; the brief remarks in this (necessarily) short section cannot and should not be taken as anything more than suggestive.

Let us consider some arguments against stochastic grammars as models of linguistic competence. In every proposal that I am aware of, the probabilities that one finds in these models are drawn from corpora. One corpus that is widely applied is derived from the *New York Times*. But no child learns English by being read to from the *Times*! Another is the ‘Switchboard Corpus’, a database of spontaneous telephone conversations by over five hundred American English speakers (Dick & Elman 2001). The Switchboard Corpus explicitly encompasses conversations from a wide variety of speech communities. But how could usage facts from a speech community to which one does not belong have any relevance whatsoever to the nature of one’s grammar? There is no way that one can draw conclusions about the grammar of an individual from usage facts about communities, particularly communities from which the individual receives no speech input.

There are many nonsequiturs in the literature that arise from ignoring this simple fact. So, Manning (2002a) observes that Pollard and Sag (1994) consider sentence 34a grammatical but star sentence 34b.

- (34) a. We regard Kim as an acceptable candidate.
 b. *We regard Kim to be an acceptable candidate.

Manning then produces examples from the *New York Times* of sentences like 34b. Perhaps 34b is generated by Pollard’s grammar and perhaps it is not. Perhaps 34b is generated by Sag’s grammar and perhaps it is not. But we will never find out by reading the *New York Times*. The point is that we do not have ‘group minds’. No input data that an individual did not experience can be relevant to the nature of his or her grammar.

But even when this methodological problem is overcome, the evidence for probabilities being associated with grammatical elements seems pretty weak. The numbers are overwhelmingly epiphenomenal. Let us examine another example from Manning 2002a.

- (35) a. It is unlikely that the company will be able to meet this year’s revenue forecasts.
 b. #That the company will be able to meet this year’s revenue forecasts is unlikely.

Manning points out that we are far more likely to say 35a than 35b and suggests that this likelihood forms part of our knowledge of grammar. No it does not. It is part of our USE of language that, for the processing reasons already discussed, speakers tend to avoid sentences with heavy subjects. As a consequence, one is more likely to say things like 35a than 35b. So there is no reason to conclude that grammars themselves reveal that likelihood.

The probability of using some grammatical element might arise as much from real-world knowledge and behavior as from parsing ease. Wasow (2002), for example, notes that we are much more likely to use the verb *walk* intransitively than transitively, as in 36.

- (36) a. Sandy walked (to the store).
 b. Sandy walked the dog.

He takes that fact as evidence that stochastic information needs to be associated with subcategorization frames. But to explain the greater frequency of sentence types like 36a than 36b, it suffices to observe that walking oneself is a more common activity than walking some other creature. It is not a fact about grammar. Furthermore, since we speak in CONTEXT, the probability that we will say whatever varies with that context. To give an example, Abney (1996) suggests that grammars have weighted probabilities for different parses of the same string. So consider phrase 37:

- (37) the shooting of the hunters

It is apparently the case that overwhelmingly in structures of the form ‘the gerund of NP’, the NP is interpreted as the object of the gerund, not as the subject. That is a raw fact about frequency of interpretation. But our interpretation of a structurally ambiguous string is determined in large part by real-word contextual factors that have nothing at all to do with grammar, no matter how broadly defined. If somebody says sentence 38 to me:

(38) The shooting of the hunters will shock the cruelty-to-animals people.

while we are looking at men with rifles standing in front of piles of bullet-ridden deer, then I would guess that the probability of the SUBJECT interpretation of the gerund would jump to 99.9%. One would still like to know why in general (that is, without special context) the object reading is more natural than the subject reading. Most likely this is so because there is an alternative means of expressing the subject reading where the preposition’s thematic role is explicit:

(39) the shooting by the hunters

It seems to me that the most natural treatment is to say that the phrase is grammatically ambiguous and that extragrammatical factors determine which reading is both statistically preferred and likely to be preferred in a given instance.

Certainly language users and hence their grammars are sensitive to frequency. But from the fact that Y is sensitive to X, it does not follow that X is part of the same system that characterizes Y. Stochastic grammar is no more defensible as an approach to language and mind than would be a theory of vision that tries to tell us what we are likely to look at.

There is a problem with stochastic approaches to syntax that does not arise with analogous approaches to phonology. (Here I am indebted to the discussion of the variable rules of sociolinguistics in Lavandera 1978.) Lavandera pointed out that the choice of phonological variants is purely a social matter. But syntactic variants often differ in the meaning that they convey. Viewed from that angle, assigning probabilities to structures or constraints seems especially problematic. The probabilities may be more a function of the meaning that one wants to convey than of some inherent property of the structure itself.

And finally there is the question of genre. Surely the probability of use of whatever construction is a function of the particular genre. But there is an indefinite number of genres. Biber (1995) has discussed forty-five different registers of spoken and written English and says that there are many more. One thinks of all the studies of the differences between the speech of men and women, gays and straights, old people and young people, and so on. Consider an example of how the multiplicity of genres cuts at the heart of stochastic grammar. Whenever one flies, one hears some bizarre (but genre-normal) syntax and lexicon from the flight crew.¹¹

- (40) a. We do request that you remain in your seats until the seat belt light is turned off.
 b. We are ready to depart the gate.
 c. Please turn off all electronic devices and stow these devices (*them).
 d. . . . until you are further notified.
 e. . . . you_i, will be free to move about the cabin as required (UNDERSTOOD: by you_i, not *by [the flight crew]_j)
 f. Takeoff will be shortly.

¹¹ I would like to thank Ralph Fasold and Jerrold Sadock for providing me with some of these examples.

- g. We hope that you will enjoy your stay in the Seattle area, or wherever your destination may take you.

What do the facts of 40 imply for our concerns? Do flight attendants have different stochastic grammars from the rest of us? Or only when they are working? Does my grammar change when I am on an airplane? Surely, it is simpler to say that part of one's training in the airline industry is in the use of otherwise ungrammatical sentence types. And those who are not airline professionals learn to interpret the meaning and appropriate use of such sentence types by appealing to their grammatical resources and world knowledge.¹²

To sum up, probabilistic information drawn from corpora is of the utmost value for many aspects of linguistic inquiry. But it is all but useless for providing insights into the grammar of any individual speaker.

11. GRAMMARS ARE NOT FRAGILE, FLUID, AND TEMPORARY. One also has to take issue with the view, expressed so often by advocates of usage-based grammar, that grammars are fragile, fluid, temporary objects. To repeat part of the Bybee and Hopper quote (2001a:2–3): '[E]mergent structures are unstable and manifested stochastically . . . From this perspective, mental representations are seen as provisional and temporary states of affairs that are sensitive, and constantly adapting themselves, to usage'. As Robert Levine pointed out to me, to read passages like that one would think normal human languages are not any different from trade pidgins like Chinook Jargon, where there are hardly any rules and communication is largely based on world-knowledge and context. In my view, one of the basic things to explain about grammars is their stability, at least where there is no significant language contact to complicate things. Consider some examples. I suspect that we could carry on a conversation with Shakespeare, who lived four hundred years ago. And the problems we would have with him would more likely be lexical and low-level phonological rather than syntactic. Preposition stranding survives, despite its being functionally odd, typologically rare, and the object of prescriptivist attack for centuries. Modern Icelanders can read the sagas from a thousand years ago. Spanish and Italian are close to being mutually intelligible, despite the fact that their common ancestor was spoken over fifteen hundred years ago and there has been no significant contact between the masses of speakers of the two languages. And so on.

12. THE EVOLUTIONARY ORIGINS OF GRAMMAR LIE IN CONCEPTUAL STRUCTURE, NOT IN COMMUNICATION. Let us change gears fairly radically and speculate on the origins of human language. I think that the most plausible evolutionary scenario might help explain why the disparity between grammar and usage is as great as it is.

When linguists started turning to questions of language origins and evolution in the early 1990s, a near consensus began to develop on a central issue. In a nutshell, the idea was that the roots of grammar lay in hominid conceptual representations and that the shaping of grammar for communicative purposes was a later development. For the most part that was taken to mean that syntax is grounded in predicate-argument structure, that is, representations embodying actors, actions, and entities acted upon, though other aspects of conceptualization were sometimes pointed to as possible antecedents of grammar. Here are some quotes from representative work, with key passages emphasized.

¹² Perhaps the more systematic aspects of one's extragrammatical knowledge are organized in the sort of 'user's manual' discussed by Zwicky (1999).

We should search for the ancestry of language not in prior systems of animal communication but in PRIOR REPRESENTATIONAL SYSTEMS. (Bickerton 1990:23)

A far better case could be made that grammar exploited mechanisms ORIGINALLY USED for the conceptualization of topology and antagonistic forces [than for motor control] (Jackendoff 1983; Pinker 1989; Talmy 1983; 1988), but that is another story. (Pinker & Bloom 1990:726)

The syntactic category system and the conceptual category system match up fairly well. In a way, the relation between the two systems serves as a partial explication of the categorial and functional properties of syntax: SYNTAX PRESUMABLY EVOLVED AS A MEANS TO EXPRESS CONCEPTUAL STRUCTURE, so it is natural to expect that some of the structural properties of concepts would be mirrored in the organization of syntax. (Jackendoff 1990:27)

The conditions for the SUBSEQUENT DEVELOPMENT of language as a medium of communication were set by the evolution of . . . the level of conceptual structure . . . A FIRST STEP toward the evolution of this system for communication was undoubtedly the linking up of individual bits of conceptual structure to individual vocalizations. (Newmeyer 1991:10)

[T]he emergent ability, driven by the evolutionary appearance of C[onceptual] S[tructure], was the capacity to acquire meaningful, symbolic, abstract units . . . IT WOULD BE APPROPRIATE TO EXPECT ADAPTATION-BASED EXPLANATIONS TO COME INTO PLAY AT A LATER STAGE, ONCE LANGUAGE CAME TO BE USED PREFERENTIALLY AS THE HUMAN COMMUNICATION SYSTEM. (Wilkins & Wakefield 1995:179)

Let me summarize the reasons for the belief that the roots of grammar lie in prehuman conceptual structure rather than in prehuman communication.¹³ First, we have learned that the conceptual abilities of the higher apes are surprisingly sophisticated (Cheney & Seyfarth 1990, Tomasello 2000, de Waal 1996). Each passing year leads to new discoveries about their capacity for problem solving, social interaction, and so on. Not human-level, but sophisticated nevertheless. Second, the COMMUNICATIVE abilities of the higher apes are remarkably primitive (Hauser 1996). There is very little calling on their conceptual structures in communicative settings. Now let us look ahead to human language. What almost all theories of grammar have in common is a tight linkage between syntactic structure and certain aspects of conceptual structure. The basic categories of reasoning—agents, entities, patients, actions, modalities, and so on—tend to be encoded as elements of grammar. This encoding is directly built into theories like cognitive grammar and construction grammar, which do not even allow for an independent level of morphosyntactic patterning. But it is true of standard generative models too. No one denies that the links between syntactic structure and whatever one might want to call it—conceptual structure/logical structure/semantic representation—are very direct.

Now, few linguists would deny that over time (possibly evolutionary time, but surely historical time) the needs of communication HAVE helped to determine the properties of grammars. This is most evident in the shaping of grammars to allow the more rapid expression of frequently used meaningful elements than of those less frequently used. So, it is auxiliaries and negative elements that tend to contract in English, not full

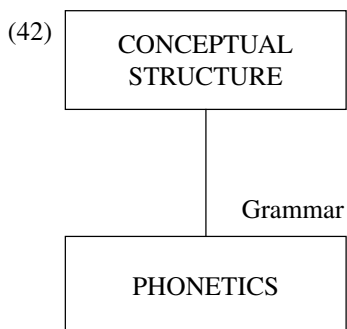
¹³ This position has been challenged recently by both functional and formal linguists. As far as the former are concerned, Charles Li has written that ‘Scholars who recognize the merit of functional linguistics agree that language emerged evolutionarily first and foremost as a vehicle of human communication, *not* as an instrument of thought. The claim of language as an instrument of thought is an intuitive and a priori claim that defies scientific verification’ (Li 2002:87). It is not clear to me why the claim that language emerged from thought is any more ‘intuitive and a priori’ than the claim that it emerged from communication. Turning to the latter, some recent work by formal linguists on language evolution rejects, at least implicitly, the idea that the evolutionary origins of language lie in conceptual structure. For Berwick (1998) all that was needed for syntax to arise was a preexisting lexicon and a single combinatorial operation called Merge. Berwick’s position is developed further in a paper co-authored by Chomsky (Hauser et al. 2002). Chomsky’s remarks in this paper represent one more example of his (to my mind, extremely troubling) retreat from advocating a richly articulated theory of Universal Grammar (for discussion, see Newmeyer 2003a).

lexical nouns and verbs. Many languages have affixes for the most commonly used concepts, such as negation, causation, comparison, and so on, but rarely for more complex infrequent concepts. Pressure for the rapid processing of spoken language has helped to shape grammars in other ways. We have seen some examples already. And one could make the same point about constraints. Universal or near-universal constraints seem designed—intuitively speaking—to ‘help’ the hearer recover pairings of fillers and gaps, antecedents and anaphors, and so on.

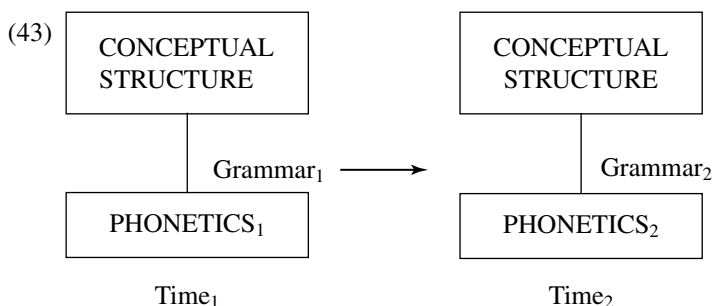
So let us consider a three-stage process in language evolution. First, there existed a level of conceptual structure:



Second came the principal evolutionary event. Conceptual structure was linked to the vocal output channel, creating for the first time a grammar that was independent of the combinatorial possibilities of conceptual structure per se, and making possible the CONVEYING of thought. That is, it made vocal communication possible. Diagram 42 illustrates.



But once grammars started to be drawn upon for real-time purposes, the constraints of real-time use began to affect their properties. In particular, grammars began to be shaped to facilitate processing, the tendency for frequently used elements to become shorter began to assert itself, and so on (43).



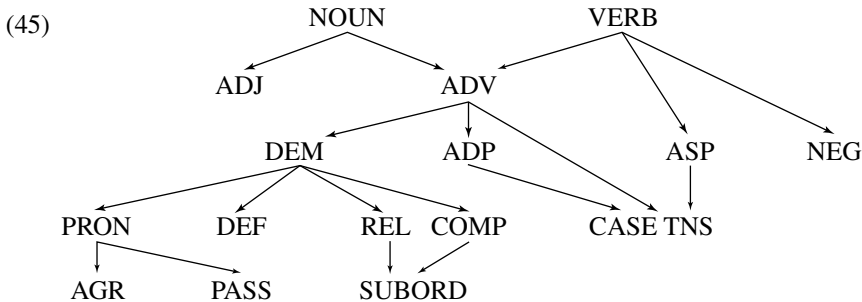
The importance of conceptual structures, that is, their antedating language per se and their forming the basis on which syntax developed, combined with the derivative appearance of language for communicative purposes, provides the evolutionary–historical basis for the disparity between grammar and usage that I have been stressing.

One of the most interesting bits of support for this model comes from discourse markers, that is, expressions like those in 44.

(44) then, I mean, y'know, like, indeed, actually, in fact, well

According to Schiffrin (1987), such expressions are essential to the makings of a coherent discourse. But as Traugott and Dasher (2002) have recently pointed out, they typically arise from something else. Usually they derive from conceptual meanings and uses constrained to the argument structure of the clause. In other words, they are central to communication, but derivative historically. That seemingly curious state of affairs loses some of its mystery if VOCAL COMMUNICATION ITSELF is derivative. Nouns and verbs trace back to nouns and verbs, because they were there from the start. The derivative nature of discourse markers points to a time when we had structured conceptual representations, but they were not coopted yet for communication.

It is not just discourse markers that seem derivative evolutionarily speaking. By and large, elements that convey grammatical meaning, and those that supply the clause with aspectual modalities, nuances of quantification and reference, and so on derive from something else and often can be traced back in their historical lineage to nouns and verbs. Heine and Kuteval (2002) have illustrated some common pathways of the development of grammatical elements as in 45.



ADP = adposition; ADJ = adjective; ADV = adverb; AGR = agreement; ASP = aspect; COMP = complementizer; DEF = definite marker; DEM = demonstrative; NEG = negation; PASS = passive; PRON = pronoun; REL = relative clause marker; SUBORD = subordination marker; TNS = tense

Heine and Kuteva draw the reasonable conclusion that in the earliest form of human language there were only two types of linguistic entities. One denoted thing-like, time-stable entities—nouns, in other words. The other denoted actions, activities, and events—verbs, in other words. Although they themselves do not go on to suggest that these facts point to a precommunicative stage for language, the conclusion seems like a reasonable one.

While I would not want to push ‘ontogeny recapitulates phylogeny’ too far, it is suggestive that discourse markers, determiners, complementizers, and so on are often relatively late to appear in children’s speech (Radford 1990). And this is the case despite the fact that the child is constantly exposed to them and they play a central role in communication.

13. CONCLUSION. The disparities between what the grammar generates and both usage BY the speaker and usefulness TO the speaker are striking. We see this in the centrality of full argument structure to grammar, even if that full argument structure is rarely expressed. We see it in the fact that speakers can make reliable judgments about sentences they would never use. And we see it in the limited place for the drive to reduce ambiguity as a functional force affecting language. The evolutionary scenario

discussed here might help to explain these facts. Prehumans possessed a rich conceptual structure, in particular one that represented predicates and their accompanying arguments. The evolutionary 'event' that underlies human language was the forging of a link between conceptual structures and the vocal output channel—in other words, the beginnings of grammar per se. But early grammar was extremely unstable. Once it was put to use for communicative purposes, it began to be shaped by those purposes. In particular, it was shaped to allow language to be produced and comprehended as rapidly as possible. But conceptual structures did not go away; they remained as part-and-parcel of every grammatical representation. They are drawn upon in every speech act, even if the actual utterance is pared down and fragmentary. At the same time, the necessarily slow rate of the transmission of human speech means that grammar has to be relieved from a big part of the job of conveying meaning in acts of language use. These considerations help explain why it makes sense to characterize grammar independently of usage. If all of this is right, then—to be sure—grammar is grammar and usage is usage.

REFERENCES

- ABNEY, STEVEN P. 1996. Statistical methods and linguistics. The balancing act: Combining symbolic and statistical approaches to language, ed. by Judith L. Klavans and Philip Resnik, 1–26. Cambridge, MA: MIT Press.
- ANDERSON, STEPHEN R., and EDWARD L. KEENAN. 1985. Deixis. Language typology and syntactic description, vol. 3: Grammatical categories and the lexicon, ed. by Timothy Shopen, 259–308. Cambridge: Cambridge University Press.
- ARNOLD, JENNIFER E.; THOMAS WASOW; ANTHONY LOSONGCO; and RYAN GINSTROM. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76.28–55.
- BERWICK, ROBERT C. 1998. Language evolution and the minimalist program: The origins of syntax. Approaches to the evolution of language: Social and cognitive bases, ed. by James R. Hurford, Michael Studdert-Kennedy, and Chris Knight, 320–40. Cambridge: Cambridge University Press.
- BIBER, DOUGLAS. 1995. Dimensions of register variation: A cross-linguistic comparison. Cambridge: Cambridge University Press.
- BICKERTON, DEREK. 1990. Language and species. Chicago: University of Chicago Press.
- BOD, RENS. 1998. Beyond grammar: An experience-based theory of language. (CSLI lecture notes 88.) Stanford, CA: CSLI Publications.
- BOD, RENS; JENNIFER HAY; and STEFANIE JANNEDY (eds.) 2003. Probabilistic linguistics. Cambridge, MA: MIT Press.
- BRESNAN, JOAN W.; SHIPRA DINGARE; and CHRISTOPHER D. MANNING. 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. Proceedings of the LFG-01 conference, ed. by Miriam Butt and Tracy H. King, 13–32. Stanford, CA: CSLI Publications.
- BYBEE, JOAN L. 1985. Morphology: A study of the relation between meaning and form. (Typological studies in language 9.) Amsterdam: John Benjamins.
- BYBEE, JOAN L., and PAUL J. HOPPER. 2001a. Introduction to frequency and the emergence of linguistic structure. Frequency and the emergence of linguistic structure, ed. by Joan L. Bybee and Paul Hopper, 1–24. Amsterdam: John Benjamins.
- BYBEE, JOAN L., and PAUL HOPPER (eds.) 2001b. Frequency and the emergence of linguistic structure. (Typological studies in language 45.) Amsterdam: John Benjamins.
- CHEN, MATTHEW, and WILLIAM S.-Y. WANG. 1975. Sound change: Actuation and implementation. *Language* 51.225–81.
- CHENEY, DOROTHY, and ROBERT SEYFARTH. 1990. How monkeys see the world. Chicago: University of Chicago Press.
- CHOMSKY, NOAM. 1975. Reflections on language. New York: Pantheon.
- CHOMSKY, NOAM. 1995. The minimalist program. Cambridge, MA: MIT Press.

- CHOMSKY, NOAM. 2000. Minimalist inquiries: The framework. Step by step: Essays on minimalist syntax in honor of Howard Lasnik, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.
- CHOMSKY, NOAM. 2001. Derivation by phrase. Ken Hale: A life in language, ed. by Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- COMRIE, BERNARD. 1989. Language universals and linguistic typology. 2nd edn. Chicago: University of Chicago Press.
- COMRIE, BERNARD. 1998. Reference-tracking: Description and explanation. *Sprachtypologie und Universalienforschung* 51.335–46.
- COWART, WAYNE. 1997. Experimental syntax: Applying objective methods to sentence judgments. Newbury Park, CA: SAGE Publications.
- CRAIG, COLETTE G. 1987. The Rama language: A text with grammatical notes. *Journal of Chibchan Studies* 5.
- CRAIN, STEPHEN, and CECILE MCKEE. 1986. Acquisition of structural restrictions on anaphora. *North Eastern Linguistic Society* 16.94–110.
- CROFT, WILLIAM. 1990. Typology and universals. Cambridge: Cambridge University Press.
- CROFT, WILLIAM. 2000. Explaining language change: An evolutionary approach. London: Longman.
- DE WAAL, FRANS B. M. 1996. Good natured: The origins of right and wrong in humans and other animals. Cambridge, MA: Harvard University Press.
- DICK, FREDERIC, and JEFFREY L. ELMAN. 2001. The frequency of major sentence types over discourse levels: A corpus analysis. *Newsletter of the Center for Research in Language, University of California, San Diego* 13.3–19.
- DRYER, MATTHEW S. 1997. On the six-way word order typology. *Studies in Language* 21.69–103.
- DU BOIS, JOHN. 1985. Competing motivations. Iconicity in syntax, ed. by John Haiman, 343–65. Amsterdam: John Benjamins.
- DU BOIS, JOHN. 1987. The discourse basis of ergativity. *Language* 63.805–55.
- ELMAN, JEFFREY L.; ELIZABETH A. BATES; MARK H. JOHNSON; ANNETTE KARMILOFF-SMITH; DOMENICO PARISI; and KIM PLUNKETT. 1997. Rethinking innateness: A connectionist perspective on development. Cambridge, MA: MIT Press.
- EMONDS, JOSEPH E. 1976. A transformational approach to English syntax. New York: Academic Press.
- ENGLAND, NORA C. 1988. Mam voice. Passive and voice, ed. by Masayoshi Shibatani, 525–45. Amsterdam: John Benjamins.
- FALTZ, LEONARD M. 1977. Reflexivization: A study in universal syntax. Berkeley: University of California dissertation. [Published, New York: Garland, 1985.]
- FOX, BARBARA A. 1994. Contextualization, indexicality, and the distributed nature of grammar. *Language Sciences* 16.1–38.
- GIDDENS, ANTHONY. 1984. The constitution of society: Outline of the theory of structuration. Cambridge: Polity Press.
- GILLIGAN, GARY M. 1987. A cross-linguistic approach to the pro-drop parameter. Los Angeles: University of Southern California dissertation.
- GIVÓN, TALMY. 1995. Functionalism and grammar. Amsterdam: John Benjamins.
- GREENBERG, JOSEPH H. 1963. Some universals of language with special reference to the order of meaningful elements. *Universals of language*, ed. by Joseph Greenberg, 73–113. Cambridge, MA: MIT Press.
- HARRIS, ROY. 1980. The language-makers. Ithaca: Cornell University Press.
- HARRIS, ROY. 1981. The language myth. London: Duckworth.
- HASPELMATH, MARTIN. 1999. Explaining article-possessor complementarity: Economic motivation in noun phrase syntax. *Language* 75.227–43.
- HAUSER, MARC D. 1996. The evolution of communication. Cambridge, MA: MIT Press.
- HAUSER, MARC D.; NOAM CHOMSKY; and W. TECUMSEH FITCH. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298.1569–79.
- HAWKINS, JOHN A. 1988. Explaining language universals. *Explaining language universals*, ed. by John A. Hawkins, 3–28. Oxford: Basil Blackwell.
- HAWKINS, JOHN A. 1994. A performance theory of order and constituency. (*Cambridge studies in linguistics* 73.) Cambridge: Cambridge University Press.

- HAWKINS, JOHN A. 2003. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- HEINE, BERND, and TANIA KUTEVA. 2002. On the evolution of grammatical forms. The transition to language, ed. by Alison Wray, 376–97. Oxford: Oxford University Press.
- HIRSH-PASEK, KATHY, and ROBERTA GOLINKOFF. 1996. *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.
- HOPPER, PAUL J. 1987. Emergent grammar. *Berkeley Linguistics Society* 13.139–57.
- HOPPER, PAUL J. 1988. Emergent grammar and the apriori grammar postulate. *Linguistics in context: Connecting observation and understanding*, ed. by Deborah Tannen, 117–34. Norwood, NJ: Ablex.
- HOPPER, PAUL J. 1998. Emergent grammar. *The new psychology of language: Cognitive and functional approaches to language structure*, ed. by Michael Tomasello, 155–75. Mahwah, NJ: Lawrence Erlbaum.
- HOPPER, PAUL J., and SANDRA A. THOMPSON. 1980. Transitivity in grammar and discourse. *Language* 56.251–99.
- HORN, LAURENCE R. 1984. Toward a new taxonomy of pragmatic inference: q- and r-based implicature. *Meaning, form, and use in context: Linguistic applications*, ed. by Deborah Schiffrin, 11–42. Washington, DC: Georgetown University Press.
- HORN, LAURENCE R. 1993. Economy and redundancy in a dualistic model of natural language. *SKY 1993: Yearbook of the Linguistic Association of Finland*, ed. by Susanna Shore and Maria Vilkuna, 33–72. Helsinki: Linguistic Association of Finland.
- JACKENDOFF, RAY. 1983. *Semantics and cognition*. Cambridge, MA: MIT Press.
- JACKENDOFF, RAY. 1990. *Semantic structures*. Cambridge, MA: MIT Press.
- JACOBSEN, WILLIAM H., JR. 1980. Inclusive/exclusive: A diffused pronominal category in native western North America. *Chicago Linguistic Society* 16.2.204–27.
- JELINEK, FRED. 1997. *Statistical methods in speech recognition*. Cambridge, MA: MIT Press.
- JOHANSSON, STIG, and KNUT HOFLAND. 1989. *Frequency analysis of English vocabulary and grammar based on the LOB corpus, vol. 1: Tag frequencies and word frequencies*. Oxford: Clarendon Press.
- JOSEPH, BRIAN D. 2001. Is there such a thing as ‘grammaticalization’? *Language Sciences* 23.2–3.163–86.
- JURAFSKY, DAN, and JAMES H. MARTIN. 2000. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.
- LABOV, WILLIAM. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45.716–62.
- LABOV, WILLIAM. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- LABOV, WILLIAM. 1994. *Principles of linguistic change, vol. 1: Internal factors*. (Language in society 20.) Oxford: Blackwell.
- LAKOFF, GEORGE. 1970/1972. Linguistics and natural logic. *The semantics of natural language*, ed. by Donald Davidson and Gilbert Harmon, 545–665. Dordrecht: Reidel.
- LAKOFF, GEORGE. 1973. Fuzzy grammar and the performance/competence terminology game. *Chicago Linguistic Society* 9.271–91.
- LAKOFF, GEORGE. 1974. Interview. *Discussing language: Dialogues with Wallace L. Chafe, Noam Chomsky, Algirdas J. Greimas [and others]*, ed. by Herman Parret, 151–78. The Hague: Mouton.
- LAKOFF, GEORGE, and MARK JOHNSON. 1999. *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.
- LAKOFF, GEORGE, and JOHN R. ROSS. 1967/1976. Is deep structure necessary? *Syntax and semantics, vol. 7: Notes from the linguistic underground*, ed. by James D. McCawley, 159–64. New York: Academic Press.
- LAMBRECHT, KNUD. 1987. On the status of SVO sentences in French discourse. *Coherence and grounding in discourse*, ed. by Russell Tomlin, 217–62. Amsterdam: John Benjamins.
- LANGACKER, RONALD W. 1987. *Foundations of cognitive grammar, vol. 1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- LANGACKER, RONALD W. 2000. A dynamic usage-based model. *Usage-based models of language*, ed. by Michael Barlow and Suzanne Kemmer, 1–63. Stanford, CA: CSLI Publications.

- LAVANDERA, BEATRIZ R. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7.171–82.
- LEVELT, WILLEM J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- LEVINSON, STEPHEN C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- LI, CHARLES N. 2002. Missing links, issues and hypotheses in the evolutionary origin of language. The evolution of language out of pre-language, ed. by Talmy Givón and Bertram F. Malle, 83–106. Amsterdam: John Benjamins.
- LIGHTFOOT, DAVID W. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.
- MACWHINNEY, BRIAN. 2000. Connectionism and language learning. Usage-based models of language, ed. by Michael Barlow and Suzanne Kemmer, 121–49. Stanford, CA: CSLI Publications.
- MANNING, CHRISTOPHER D. 2002a. Probabilistic syntax. Probabilistic linguistics, ed. by Rens Bod, Jennifer Hay, and Stefanie Jannedy, 289–341. Cambridge, MA: MIT Press.
- MANNING, CHRISTOPHER D. 2002b. Review of *Beyond grammar: An experience-based theory of language*, by Rens Bod. *Journal of Linguistics* 38.441–42.
- MANNING, CHRISTOPHER D., and HINRICH SCHÜTZE. 1999. *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- MARTIN, WILLIAM A.; KENNETH W. CHURCH; and RAMESH S. PATEL. 1987. Preliminary analysis of the breadth-first parsing algorithm: Theoretical and experimental results. *Natural language parsing systems*, ed. by Leonard Bolc, 267–328. Berlin: Springer Verlag.
- MCCAWLEY, JAMES D. 1968. Lexical insertion in a transformational grammar without deep structure. *Chicago Linguistic Society* 4.71–80.
- MCDANIEL, DANA, and WAYNE COWART. 1999. Experimental evidence for a minimalist account of English resumptive pronouns. *Cognition* 70.B15–B24.
- MITHUN, MARIANNE. 1987. Is basic word order universal? Coherence and grounding in discourse, ed. by Russ Tomlin, 281–328. Amsterdam: John Benjamins.
- MORGAN, JERRY L. 1973. Sentence fragments and the notion 'sentence'. *Issues in linguistics: Papers in honor of Henry and Renée Kahane*, ed. by Braj B. Kachru, Robert B. Lees, Yakov Malkiel, Angelina Pietrangeli, and Sol Saporta, 719–51. Urbana: University of Illinois Press.
- NANNI, DEBORAH L. 1978. The 'easy' class of adjectives in English. Amherst: University of Massachusetts dissertation.
- NEUMEYER, FREDERICK J. 1986. *Linguistic theory in America*. 2nd edn. New York: Academic Press.
- NEUMEYER, FREDERICK J. 1991. Functional explanation in linguistics and the origins of language. *Language and Communication* 11.3–28.
- NEUMEYER, FREDERICK J. 1998. *Language form and language function*. Cambridge, MA: MIT Press.
- NEUMEYER, FREDERICK J. 2001. Where is functional explanation? *Chicago Linguistic Society* 37.2.99–122.
- NEUMEYER, FREDERICK J. 2003a. Review of *On nature and language*, by Noam Chomsky; *The language organ: Linguistics as cognitive physiology*, by Stephen R. Anderson and David W. Lightfoot; *Language in a Darwinian perspective*, by Bernard H. Bichakjian. *Language* 79.583–99.
- NEUMEYER, FREDERICK J. 2003b. 'Basic word order' in formal and functional linguistics and the typological status of 'canonical' sentence types. *Contrastive analysis in language: Identifying linguistic units of comparison*, ed. by Dominique Willems, Bart Defrancq, Timothy Coleman, and Dirk Noël. Basingstoke: Palgrave to appear.
- NICHOLS, JOHANNA. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- PAYNE, DORIS L. 1990. *The pragmatics of word order: Typological dimensions of verb-initial languages*. Berlin: Mouton de Gruyter.
- PAYNE, DORIS L. 1992. Nonidentifiable information and pragmatic order rules in 'O'dham. *Pragmatics of word order flexibility*, ed. by Doris L. Payne, 137–66. Amsterdam: John Benjamins.

- PHILLIPS, BETTY S. 1984. Word frequency and the actuation of sound change. *Language* 60.320–42.
- PINKER, STEVEN. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
- PINKER, STEVEN, and PAUL BLOOM. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13.707–84.
- POLLARD, CARL, and IVAN A. SAG. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
- POPE, EMILY. 1971. Answers to yes-no questions. *Linguistic Inquiry* 2.69–82.
- RADFORD, ANDREW. 1990. *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English*. Oxford: Blackwell.
- ROEPER, THOMAS, and MUFFY SIEGEL. 1978. A lexical transformation for verbal compounds. *Linguistic Inquiry* 9.199–260.
- ROSS, JOHN R. 1973a. A fake NP squish. *New ways of analyzing variation in English*, ed. by Charles-James N. Bailey and Roger W. Shuy, 96–140. Washington, DC: Georgetown University Press.
- ROSS, JOHN R. 1973b. Nouniness. *Three dimensions of linguistic theory*, ed. by Osamu Fujimura, 137–258. Tokyo: TEC Company, Ltd.
- ROSS, JOHN R. 1973c. The penthouse principle and the order of constituents. *You take the high node and I'll take the low node*, ed. by Claudia Corum, C. Cedric Smith-Stark, and Ann Weiser, 397–422. Chicago: Chicago Linguistic Society.
- RUMELHART, DAVID E., and JOHN L. MCCLELLAND. 1986. On learning the past tenses of English verbs. *Parallel distributed processing: Explorations in the microstructure of cognition*, ed. by John L. McClelland and David E. Rumelhart, 216–71. Cambridge, MA: MIT Press.
- SCANCARELLI, JANINE S. 1985. Referential strategies in Chamorro narratives: Preferred clause structure and ergativity. *Studies in Language* 9.335–62.
- SCHIFFRIN, DEBORAH. 1987. *Discourse markers*. (Studies in interactional sociolinguistics 5.) Cambridge: Cambridge University Press.
- SCHUCHARDT, HUGO. 1885/1972. On sound laws: Against the Neogrammarians. *Schuchardt, the Neogrammarians, and the transformational theory of phonological change*, ed. by Theo Vennemann and Terence Wilbur, 39–72. Frankfurt: Athenäum.
- SCHUETZE-COBURN, STEPHAN. 1987. *Topic management and the lexicon: A discourse profile of three-argument verbs in German*. Los Angeles: University of California, Los Angeles M.A. thesis.
- SCHÜTZE, CARSON. 1996. *The empirical basis of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.
- SEIDENBERG, MARK S. 1997. Language acquisition and use: Learning and applying probabilistic constraints. *Science* 275.1599–603.
- SMITH, WENDY. 1996. Spoken narrative and preferred clause structure: Evidence from Modern Hebrew discourse. *Studies in Language* 20.163–89.
- SMOLENSKY, PAUL. 1999. Grammar-based connectionist approaches to language. *Cognitive Science* 23.589–613.
- STEEDMAN, MARK. 1999. Connectionist sentence processing in perspective. *Cognitive Science* 23.615–34.
- TALMY, LEONARD. 1983. How language structures space. *Spatial orientation: Theory, research, and application*, ed. by Herbert L. Pick and Linda P. Acredolo, 225–82. New York: Plenum Press.
- TALMY, LEONARD. 1985/2000. Lexicalization patterns: Semantic structure in lexical forms. *Toward a cognitive semantics*, ed. by Leonard Talmy, 21–212. Cambridge, MA: MIT Press.
- TALMY, LEONARD. 1988. Force dynamics in language and cognition. *Cognitive Science* 12.49–100.
- THOMPSON, SANDRA A., and PAUL J. HOPPER. 2001. Transitivity, clause structure, and argument structure: Evidence from conversation. *Frequency and the emergence of linguistic structure*, ed. by Joan L. Bybee and Paul Hopper, 27–60. Amsterdam: John Benjamins.
- TOMASELLO, MICHAEL (ed.) 2000. *Primate cognition (special issue)*. *Cognitive Science* 24.3.
- TRAUGOTT, ELIZABETH C., and RICHARD B. DASHER. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.

- WANG, WILLIAM S.-Y. 1969. Competing changes as a cause of residue. *Language* 45.9–25.
- WANG, WILLIAM S.-Y. (ed.) 1977. *The lexicon in phonological change*. The Hague: Mouton.
- WASOW, THOMAS. 2002. *Postverbal behavior*. Stanford, CA: CSLI Publications.
- WEBER, DAVID J. 1989. *A grammar of Huallaga (Huánaco) Quechua*. Berkeley: University of California Press.
- WILKINS, WENDY K., and JENNIE WAKEFIELD. 1995. Brain evolution and neurolinguistic preconditions. *Behavioral and Brain Sciences* 18.161–226.
- ZIPF, GEORGE. 1935. *The psychobiology of language*. New York: Houghton Mifflin.
- ZIPF, GEORGE. 1949. *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley.
- ZWICKY, ARNOLD M. 1999. *The grammar and the user's manual*. Paper presented at Linguistic Society of America Summer Institute Forum Lecture, Urbana, IL.

Department of Linguistics
University of Washington
Seattle, WA 98195-4340
[fjn@u.washington.edu]

[Received 17 January 2003;
accepted 19 June 2003]