

**From usage to grammar:
The mind's response to repetition**

Joan Bybee
University of New Mexico

MSC03-2130, Linguistics,
1 University of New Mexico,
Albuquerque, NM 87131

jbybee@unm.edu

Abstract

A usage-based view takes grammar to be the cognitive organization of one's experience with language. Aspects of that experience, for instance, the frequency of use of certain constructions or particular instances of constructions, have an impact on representation that are evidenced in speaker knowledge of conventionalized phrases, and in language variation and change. It is shown that particular instances of construction can acquire their own pragmatic, semantic and phonological characteristics. In addition, it is argued that high frequency instances of constructions undergo grammaticization processes (which produce further change), function as the central members of categories formed by constructions and retain their old forms longer under the pressure of newer formations. An exemplar model that accommodates both phonological and semantic representation is elaborated to describe the data considered.*

* I am grateful to Sandra Thompson and Rena Torres-Cacoullas for many discussions of the phenomena treated in this paper. In addition, the questions and comments after the Presidential Address in Oakland in 2005 from Ray Jackendoff, Mark Baker, Larry Horn and Janet Pierrehumbert stimulated improvements in the paper.

1. Usage-based grammar.

The observance of a separation between the use of language and its internalized structure can be traced back to Saussure's well-known distinction between langue and parole, which was adhered to by American structuralists and made its way into generative grammar via Chomsky's distinction between competence and performance. In American structuralism and in generative grammar, the goal of studying langue or competence was given highest priority and the study of language use in context has been considered not relevant to the understanding of grammar. However, other goals for linguistic research which do not isolate the study of language structure from language use have been pursued through the last few decades by a number of researchers dubbed as 'functionalists' (for instance, Greenberg 1966, Givon 1979, Hopper and Thompson 1980, Bybee 1985) with cognitive linguists more recently joining the ranks, to create a broad research paradigm under the heading of Usage-Based Theory (Barlow and Kemmer 2000, Langacker 2000, Bybee 2001).

While all linguists are likely to agree that grammar is the cognitive organization of language, a usage-based theorist would make the more specific proposal that grammar is the cognitive organization of one's experience with language. As will be shown in this paper, certain facets of linguistic experience, such as, the frequency of use of particular instances of constructions, has an impact on representation that we can see evidenced in various ways, for instance, in speakers' recognition of what is conventionalized and what is not, and even more strikingly in the nature of language change. The proposal presented here is that the general cognitive capabilities of the human brain, which allow it to categorize and sort for identity, similarity and difference, go to work on the language events a person encounters, categorizing and entering in memory these experiences. The result is a cognitive representation that can be called a grammar. This grammar, while it may be abstract, since all cognitive categories are, is strongly tied to the experience that a speaker has had with language.

In addition to presenting evidence that specific usage events affect representation, I will also address the issue of the type of cognitive representation that is necessary to accommodate the facts that are brought to light in this usage-based perspective. I will be arguing for morphosyntax, as I have for phonology, that one needs an exemplar representation for language experience, and that constructions provide an appropriate vehicle for this type of representation.

2. Converging trends in linguistic theory.

In recent years many researchers have moved towards a consideration of the effect that usage might have on representation. One practice that unites many of these researchers is a methodological one: it is common now to address theoretical issues through the examination of bodies of naturally-occurring language use. This practice has been in place for decades in the work of those who examine the use of grammar in discourse with an eye towards determining how discourse use shapes grammar, notably Givon, Thompson, Hopper, Chafe and DuBois, (e.g. Givon 1979, Hopper and Thompson 1980, Ono et al. 2000, Thompson and Hopper 2001). In addition, researchers in sociolinguistic variation, such as Labov, Sankoff and Poplack (e.g. Poplack 2001, Poplack and Tagliamonte 1999, 2001) have always relied on natural discourse to study the inherent variation in language use.

The importance of usage is also emphasized in the work on grammaticization, e.g. Bybee 2003a, b, Hopper and Traugott 1993, Poplack and Tagliamonte 1999. In fact, the study of grammaticization has played a central role in emphasizing the point that both grammatical meaning and grammatical form come into being through repeated instances of language use. This line of research along with the discourse research mentioned above indeed seeks to explain the nature of grammar through an examination of how grammar is created over time, thus setting a higher goal for linguistic explanation than that held in more synchronically-oriented theory, which requires only that an explanatory theory provide the means for adequate synchronic description (Chomsky 1957).

Of course one major impetus for the shift to analysis of natural language use is the recent availability of large electronic corpora and means for accessing particular items and patterns in such corpora. Through the work of corpus linguists, such as John Sinclair (1991), computational linguists, such as Dan Jurafsky and colleagues (e.g. Jurafsky et al. 2001, Gregory et al. 1999), those who are proposing probabilistic or stochastic grammar, such as Janet Pierrehumbert (e.g. 2001), Rens Bod (1998), access to the nature and range of experience an average speaker has with language is now within our grasp.¹ Studies of words, phrases, and constructions in such large corpora present a varying topography of distribution and frequency that can be quite different from what our intuitions have suggested. In addition, the use of large corpora for phonetic analysis provide a better understanding of the role of token frequency and specific words and collocations in phonetic variation.

At the same time a compatible view of language acquisition has been developing. The uneven distribution of words and constructions in speech to children is mirrored somewhat in the course of acquisition: children often produce their first instances of grammatical constructions only in the context of specific lexical items and later generalize them to other lexical items, leading eventually to productivity; see Tomasello, Lieven, and their colleagues (e.g. Lieven et al. 2003, Tomasello 2003, Savage et al. 2003).

3. Findings.

As linguists turn their attention to natural language use they find a fascinating new source of insights about language. One finding that seems to hold across many studies and has captured the interest of researchers is that both written and spoken discourse are characterized by the high use of conventionalized word sequences, which include sequences which we might call formulaic language and idioms, but also conventionalized collocations (sometimes called 'prefabs'). Idioms are conventionalized word sequences that usually contain ordinary words and predictable morpho-syntax, but have extended meaning (usually of a metaphorical nature), as in these examples: pull strings, level playing field, too many irons in the fire. Idioms are acknowledged to need lexical representation because of the unpredictable aspects of their meaning, but as Nunberg et al. (1994) point out, they are not completely isolated since many aspects of their meaning and form derive from more general constructions and the meaning of the component words in other contexts. Idioms provide evidence for organized storage in which sequences of words can have lexical representation while still being associated with other occurrences of the same words, as in this diagram from Bybee 1998.²

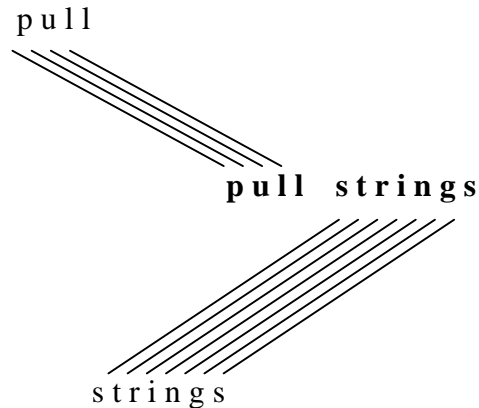


Figure 1. The relation of an idiom to its lexical components.

Idioms have a venerable history in linguistic study, but prefabs or collocations have attracted less attention through recent decades (but see Bolinger 1961, Pawley and Syder 1983, Sinclair 1991, Wray 2002, Biber et al. 1999, and Erman and Warren 2000). Prefabs are word sequences that are conventionalized, but predictable in all other ways. For example, word sequences such as prominent role, mixed message, beyond repair, to need help. In addition, phrasal verbs (finish up, burn down) and verb-preposition pairings (interested in, think of, think about), which are pervasive in English as well as other languages, can be considered prefabs. These conventional collocations occur repeatedly in discourse and are known to represent the conventional way of expressing certain notions (Erman and Warren 2000, Sinclair 1991, Wray 2002). Erman and Warren 2000 found that what they call prefabricated word combinations constitute about 55% of both spoken and written discourse. Speakers recognize prefabs as familiar, which indicates that these sequences of words must have memory storage despite being largely predictable in form and meaning.

The line between idiom and prefab is not always clear since many prefabs require a metaphorical stretch for their interpretation. The following may be intermediate examples, where at least one of the words requires a more abstract interpretation: break a habit, change hands, take charge of, give (s.o.) plenty of time, drive (s.o.) crazy. I bring up these intermediate cases to demonstrate the gradient nature of these phenomena; the lack of a clear boundary between idioms and prefabs would also suggest that both types of expression need memory storage.

What we see instantiated in language use is not so much abstract structures as specific instances of such structure that are used and re-used to create novel utterances. This point has led Hopper (1987) to propose grammar as emergent from experience, mutable, and ever coming into being rather than static, categorical, and fixed. Viewed in this way, language is a complex dynamic system similar to complex systems that have been identified, for instance, in biology (Lindblom et al. 1984, Larsen-Freeman 1997). It does not have structure a priori, but rather the apparent structure emerges from the repetition of many local events (in this case speech events). The present paper describes some data that help us understand what some of the properties of an emergent, usage-based grammar might be.

4. Goals of the article.

There are a number of important consequences of the fact that speakers are familiar with certain multi-word units. For the present paper I will focus on the implications of the fact that the use of language is lexically particular; certain words tend to be used in certain collocations or constructions. My goal will be to explore the implications of this fact for representation. I will discuss a series of cases in which there is evidence that lexically particular instances of constructions or word sequences are stored in memory and accessed as a unit. I will further discuss facts that show that the frequency of use of such lexically particular collocations must also be registered in representation because they are instrumental in certain types of change. I will argue that in order to represent the facts of usage, as well as the facts of change that eventually emerge from this usage, we need to conceive of grammar as based on constructions and as having an exemplar representation in which specific instances of use affect representation. The model to be proposed, then, uses a type of exemplar representation with constructions as the basic unit of morphosyntax (see sections 6-9 and 13).

Five types of evidence will be discussed. In section 10 the importance of frequency in the developing autonomy of new constructions in grammaticization is presented. In the following section, I discuss the effects of context and frequency of use on the development of conventionalized collocations and grammatical constructions. Section 12 briefly treats phonological reduction in high frequency phrases. Section 14 turns to the organization of categories within constructions where it is seen that in some cases high frequency exemplars serve as the central members of categories. Finally, the fact that high frequency exemplars of constructions can resist change is taken as evidence that such exemplars have cognitive representation.

5. Frequency effects on processing and storage.

Before turning to this evidence, I will briefly review three effects of token frequency that have been established in recent literature:

First, high frequency words and phrases undergo phonetic reduction at a faster rate than low and mid frequency sequences (Schuchardt 1885, Fidelholtz 1975, Hooper 1976, Bybee and Scheibman 1999, Bybee 2000b, 2001). This Reducing Effect applies to phrases of extreme high frequency such as I don't know which shows the highest rate of don't reduction (Bybee and Scheibman 1999), and also to words of all frequency levels undergoing gradual sound change, such as English final t/d deletion or Spanish [ð]deletion, both of which affect high frequency words earlier than low frequency words (Bybee 2001, 2002, Gregory et al. 1999). The explanation for this effect is that the articulatory representation of words and sequences of words are made up of neuromotor routines. When sequences of neuromotor routines are repeated, their execution becomes more fluent. This increased fluency is the result of representing the repeated sequence at a higher level as a single unit (Anderson 1993, Boyland 1996).

A second effect of token frequency (the Conserving Effect) relates to the morphosyntactic structure of a string. High frequency sequences become more entrenched in their morpho-syntactic structure and resist change on the basis of more productive patterns. Thus among English irregular verbs the low frequency verbs are more likely to regularize (weep, weaped) while the high frequency verbs maintain their irregularity (keep, kept). My proposal to explain this tendency (Hooper 1976, Bybee

1985) is that frequency strengthens the memory representations of words or phrases making them easier to access whole and thus less likely to be subject to analogical reformation. This effect applies to syntactic sequences as well, allowing higher frequency exemplars to maintain a more conservative structure (Bybee and Thompson 1997). In Section 15 the example of the maintenance of the older type of negation in English (*no-negation*) with high frequency constructions is discussed.

The third effect (Autonomy) is related to the second one. Autonomy refers to the fact that morphologically complex forms (or strings of words) of high frequency can lose their internal structure as they become autonomous from etymologically related forms (Bybee 1985). This can be seen for example in the way that words with derivational affixes become less transparently related to their base forms as they become more frequent (Bybee, 1985; Hay 2001). Hay 2001 argues that the semantic opacity of words such as dislocate, etc. is due to the fact that their complex forms are more frequent than the bases from which they were originally derived. The effect applies to inflection only in cases of extreme high frequency where case it leads to suppletion. Thus went was formerly the past tense of wend but (for unknown reasons) it increased in frequency and moved away from wend, joining go to become the past tense of that verb. This effect also applies in grammaticization when sequences that are originally complex (such as be going to) lose their semantic and syntactic transparency and move away from other instances of the words be, go, and to.

In discussing these effects here and elsewhere, I refer to high and low frequency and to extreme high frequency without specifying exactly what these values mean in numerical terms. Thus, the Conserving Effect applies to high frequency items but Autonomy appears to affect only strings of extreme high frequency. The reducing effect appears to be graded in that the higher the frequency of the string, the greater its reduction. The phenomenon discussed in section 14 in which higher frequency items for the centers of categories requires that the item not be so high in frequency as to be autonomous. The impossibility at the moment of specifying ranges for extreme high, medium and low is only a function of the state of our knowledge. As more empirical studies appear, absolute frequency ranges for each phenomena will eventually be specifiable.

6. Construction-based representations.

For the phenomena to be examined here, representation based on constructions turns out to be highly effective. Several versions of grammar in terms of constructions have been discussed in the literature. Proposals have been made by Fillmore and Kay (e.g. Kay et al. forthcoming, Fillmore et al. 1988, Fillmore and Kay 1994), Goldberg (1995, 2003), Lakoff (1987), Langacker (1987) and Croft (2001). All of these proposals agree on the basic point that:

Cognitive representations of grammar are organized into constructions which are partially schematic, conventionalized sequences of morphemes with a direct semantic representation (Goldberg 2003).

According to Goldberg, all of the following constitute constructions:

- (1) Idioms with fixed lexical content: go great guns
- (2) Idioms that are partially filled: jog <someone's> memory
- (3) Constructions with some fixed material: he made his way through the crowd
- (4) Fully abstract constructions: they gave him an award

It is interesting to note that almost all constructions contain some explicit morphological material, tying them fairly concretely to specific words or morphemes (e.g. way and the possessive pronoun in 3). The ditransitive construction in 4 contains no specific phonological material that identifies it as the ditransitive. Only the word order signals this. However, it should be noted that only a small class of verbs can occur in this construction so that it also has a grounding in lexical items.

In fact, the continuum in 1 through 4 shows examples from the most lexically explicit to the most schematic. Prefabs can also be considered as instances of constructions that are lexically filled. Given the high use of prefabs and idioms in natural speech, it appears that a good deal of production (and perception) refers to sequences of pre-specified lexical choices rather than to abstract grammar. For this reason, a model that builds a grammar from specific instances of language use, such as an exemplar model or a connectionist model, seems appropriate. For present purposes, I focus on representation in an exemplar model.

7. Exemplar representation.

Several versions of exemplar theory have been proposed in the psychology literature on categorization (Nosofsky 1988, Goldinger 1996). The version of exemplar theory to be adopted here has found its way into linguistics as a means of representing phonetic variation (K. Johnson 1997, Pierrehumbert 2001, 2002). In this model, every token of experience is classified and placed in a vast organizational network as a part of the decoding process. The major idea behind exemplar theory is that this matching process has an effect on the representations themselves; new tokens of experience are not decoded and then discarded, but rather they impact memory representations. In particular, a token of linguistic experience that is identical to an existing exemplar is mapped onto that exemplar, strengthening it. Tokens that are similar but not identical (differing in slight ways in meaning, phonetic shape, pragmatics) to existing exemplars are represented as exemplars themselves and are stored near similar exemplars to constitute clusters or categories. Thus the phonetic shape of a word might consist of a set of phonetic exemplars that are very similar to one another.³

Exemplar clusters can also be arranged hierarchically. A set of exemplars that are judged to be similar phonetically and represent the same meaning are clustered together and are represented at a higher level as a word or phrase. Constructions emerge when phrases that bear some formal similarity as well as some semantic coherence are stored close to one another.

According to Johnson 1997, phonetic exemplars will be tagged for an array of information about their occurrence: phonetic context, semantic and pragmatic information, other linguistic context, and social context. Here I am expanding this theory to include exemplar representations of all aspects of a linguistic sign. Since words and constructions pair form with meaning, representations must contain exemplar clusters for

the meanings (semantic and pragmatic) as well as for the phonetic representations. Section 13 expands on this proposal in more detail.

In applying this kind of model to linguistic data, it is important to bear in mind just which aspects of the model are helpful to our understanding of language. Thus I focus on the following factors because they help us understand how constructions come into being and change over time.

- a. Exemplar representations allow specific information about instances of use to be retained in representation
- b. Exemplar representation provides a natural way to allow frequency of use to determine the strength of exemplars
- c. Exemplar clusters are categories that exhibit prototype effects. They have members that are more or less central to the category, rather than categorical features.

At first it might seem rather implausible to suppose that every token of language use encountered by a speaker/hearer is registered in memory. Therefore it is important to consider how these notions can be applied to language.

First, we have to move beyond the goal of structural and generative linguistics to try to establish which features or forms are stored in the lexicon and which are not (cf. Langacker's rule/list fallacy [1987]) and take a more probabilistic view of representation. A phrase that is experienced only once by an adult is likely to have only a minute impact on representation compared to all the accumulated exemplars already existing. (Compare this to a young child whose experience is much more limited: each new token of experience will have a larger impact on his/her representations.) In addition, given a highly organized network of morphemes, words, phrases, and constructions, it will be difficult to distinguish between specific storage (storing the relatively low frequency phrase such as beige curtains as a unit) and distributed storage (mapping the two words onto existing exemplars of these words) because both types of processing are occurring at the same time. Only when a sequence is repeated will access to it as a unit rather than by its parts become more efficient (Boyland 1996, Haiman 1994, Hay 2001). Thus the question of storage or non-storage will always be a probabilistic one, based on the experience of the language user.

Second, human memory capacity is quite large. Non-linguistic memories are detailed and extensive suggesting a strong memory capacity for experienced phenomena (Goldinger 1996, K. Johnson 1997). In particular it is interesting to consider how memory for repeated experiences is represented because an important feature of linguistic experience is the regular repetition of phonological strings, words, and constructions. Consider a repeated experience, such as walking from your office on campus to your classroom. You probably take much the same route each time, and although many of the perceptual details are not important, you register them anyway. Was there a pigeon on the path? Did you see your colleague coming from the other direction? These details are registered in memory so that if they are repeated, it is noted and eventually you may come to expect to see a pigeon on the path or your colleague returning from his class. At the same time, memories decay over time. If in fact you never see your colleague again on that path, you may forget the one time you saw him/her there. Both of these properties

of memory—the build up of strength in repeated memories and the loss of non-repeated memories—are important for explaining linguistic phenomena.

Third, it is clear that linguistic memories represented as exemplars can undergo considerable reorganization particularly when change is ongoing in a language. Examples of reorganization of constructions are presented in sections 10 – 12, but for now we mention a phonetic example. Consider the case of a sound change occurring at a word boundary which sets up variation in a word, such as in the final syllable of a Spanish word ending in [s] where it is reduced to [h] before a consonant but maintained before a vowel. In such cases, the phonetically conditioned variation is not maintained. Rather, the more frequent variant, the one used before a consonant is extended to use before a vowel. This gives evidence for considerable reorganization within the exemplar cluster resulting in a smaller range of variation (Bybee 2000a, 2001). (For experiments that model category formation from exemplars, see Pierrehumbert 2001, 2002; Wedel 2004.)

The cases I will discuss here provide evidence that linguistics needs a model that allows particular instances of use to affect representation. It is important to note that there are a variety of models that have been computationally tested that have this property; for instance, connectionist models work on the principle that detailed information is the basis of more general patterns. Such models have been tested on a variety of cognitive phenomena, but it may be that language will present both the toughest testing ground for such theories, as well as the best source of data for our understanding of categorization and memory, especially memory for repeated events.

8. Exemplar representation of constructions.

As mentioned above, exemplars of words or phrases that are similar on different dimensions are grouped together in cognitive storage. From such groupings a construction can emerge. For instance, an exemplar representation of a partially filled construction would have experienced tokens mapping onto the constant parts of the construction exactly, strengthening these parts, while the open slots would not match exactly. If there are similarities (in particular semantic similarities) among the items occurring in the open slot, a category for these items would begin to develop. Thus in a Spanish example that we will discuss further in section 14, adjectives following the verb quedarse ‘to become’, such as quieto, tranquilo, inmóvil, would be categorized together due to their semantic coherence and this category would then predict novel uses of quedarse + adjective (Bybee and Eddington, to appear).

(5) Example: Spanish quedarse + ADJECTIVE ‘become ADJ’

<u>quedarse</u>	$\left\{ \begin{array}{l} \underline{\text{tranquilo}} \\ \underline{\text{quieto}} \\ \underline{\text{inmóvil}} \end{array} \right\}$	$\left. \begin{array}{l} \text{‘tranquil’} \\ \text{‘quiet, still’} \\ \text{‘still, immobile’} \end{array} \right\}$
-----------------	---	---

Further details about the mapping of form and meaning in constructions will be presented in section 13.

9. Effects of repetition on particular instances of constructions.

The next three sections of this article will be a discussion of the effects of repetition and context of use on particular exemplars of constructions. The examples provided are intended as empirical evidence that specific instances (exemplars) of constructions are part of the cognitive representation of language and that frequency of use has an impact of the nature of these representations. There are various degrees of effect, depending upon the extent of the frequency.

1. low levels of repetition lead to conventionalization only (as in prefabs and idioms)
2. higher levels of repetition can lead to the establishment of a new construction with its own categories
3. extreme high frequency leads to the grammaticization of the new construction and the creation of grammatical morphemes and changes in constituency.

The discussion begins with the grammaticization (listed third) because, as a phenomenon of linguistic change, its properties have been well studied.

10. Grammaticization. The idea that frequency of use affects the formation of grammar has been explored in some detail in studies on grammaticization (Haiman 1994, Boyland 1996, Bybee 2003). Grammaticization (or grammaticalization) is the creation of a new grammatical morpheme and a new construction out of a particular instance of an old construction (Heine et al. 1991, Hopper and Traugott 1993, Bybee et al. 1994). In particular, an existing construction with specific lexical items in it becomes more frequent, changes in various ways, and becomes a new construction. By examining how this process occurs we learn about the cognitive representation of constructions. The example to be discussed is the development of be going to as a marker first of intention and then of future in English (Danchev and Kytö 1994).

In the mid to late 16th century, English be going to was an exemplar of a general purpose construction sometimes used in the progressive, as shown in these examples from the works of William Shakespeare:

- (6) Don Alphonso,
 With other gentlemen of good esteem,
Are journeying to salute the emperor
 And to commend their service to his will. (Two Gentlemen of Verona I.3)
- (7) I was sending to use Lord Timon myself... (Timons of Athens II.2)

The most frequent instance of this construction used go as the main verb. In the complete works of William Shakespeare there six instances of be going to in a purpose construction and only two instances with other verbs (examples 6 and 7).⁴ This construction gradually gains in frequency over subsequent centuries (Danchev and Kytö 1994).

- (8) ...the kings
and the princes, our kindred, are going to see the queen's picture.
(Winter's Tale V.2)
- (9) Sir, the Germans desire to have three of your
horses: the duke himself will be to-morrow at
court, and they are going to meet him. (Merry Wives of Windsor IV.3)

Exemplars of the purpose clause construction are originally grouped together.



As the sequence with the originally lexical verb go occurs more frequently, it will gradually gain strength (represented in 10 by its larger font size). There are several consequences of the strength of representation afforded by high frequency in experience, as will be discussed below in greater detail. The general claim is that this more frequent sequence gradually moved away from its source construction becoming more autonomous. Thus a new construction was created and as we shall see below, certain pragmatic inferences came to be associated with it.

- (11) [SUBJECT + be going + to + VP] INTENTION, FUTURE

As is well-known, phonological reduction affects this construction leading to the reduction of going to to [gənə] often spelled gonna.

As I argued in Bybee 2003a, the changes that take place in grammaticization are conditioned at least in part by high frequency of use. The following is a brief explanation of how frequency of use helps to condition the changes that took place in this construction. Note that all of these changes are intricately interrelated.

First, as I have already argued, phonological reduction takes place when words and phrases are often repeated (the Reducing Effect). Thus the increasing token frequency of be going to leads to the creation of a neuromotor routine that is processed as a single unit and can undergo phonological reduction defined as a decrease in the magnitude of gestures and increase in their overlap. Indeed, the highest frequency expression involving be going to is I'm going to, which is often produced as [aimənə].

Second, the autonomy of a new construction is conditioned by frequency as explained in section 5. That is, as a particular string grows more frequent, it comes to be processed as a unit rather than by its individual parts. As it is accessed more and more as a unit, it grows autonomous from the construction that originally gave rise to it. It loses its association with the purpose construction and also with the other instances of the verb go.

Third, the loss of the specific meaning of movement in space and addition of inferential meaning from the context also relies on frequency of use. The be going to construction in many contexts carried the pragmatic inference of intention, as shown in the following exchange from Two Gentlemen of Verona as cited in Hopper and Traugott 1993.

- (12) Duke Sir Valentine, whither away so fast?
Val. Please it your grace, there is a messenger
That stays in to bear my letters to my friends,
And I am going to deliver them.
(1595, Shakespeare, Two Gentlemen of Verona III.i.51)

In this example, the Duke's literal question is 'where are you going'. Valentine's answer does not specify location, but rather intention. Interestingly, that is actually what the Duke wanted to know. The inference of intention often accompanies the use of this construction. Repeated instances such as this one make 'intention' part of the meaning of the construction. As we will see in section 13, the meaning and contextual implications of a construction form an exemplar cluster much as the phonetic variants do. These clusters are susceptible to the same sort of reorganization we have discussed with respect to phonetics: high frequency semantic/pragmatic exemplars come to dominate the cluster and lower frequency exemplars may be lost, bringing about gradual semantic change.

Finally, because items that are used together frequently come to be processed together as a unit, changes in constituency and category take place. Thus going to as the constant part of this construction becomes a single unit not just phonologically, but also syntactically. As the construction acquires new nuances of meaning and loses its motion sense, the following verb is taken to be the main verb. This process, known as 'reanalysis', is viewed in a usage-based perspective as being gradual, that is, as consisting of a gradual change in the exemplar cluster (Haspelmath 1998).

Thus grammaticization, which is the major vehicle for the creation of new grammatical morphemes, demonstrates the need for the cognitive representation of instances of constructions, because if specific instances of the purpose construction (with be going to) were not registered in memory, it could not begin the process of grammaticization. In addition, the various effects that repetition have on an exemplar of a construction demonstrate that cognitive representations are affected by frequency or repetition.

11. New constructions without grammaticization.

In Bybee 2003a, I defined grammaticization as the creation of a new construction out of a particular instance of an existing construction (see also Traugott 2003). This definition is not correct because new constructions can arise without the hallmark of grammaticization, which is the creation of a new grammatical morpheme. In this section we will examine the cases of new constructions arising without grammaticization. The goal of this section is to argue that particular instances of constructions that have been experienced by a speaker must be present in cognitive representation.

It has already been noted that idioms with specific metaphorical meaning must have cognitive representation even though they are still related to the lexemes and

construction from which they arose (Nunberg et al. 1994). Similarly, as noted earlier, prefabs, which are conventionalized but more transparent in meaning, must also have a representation (because they are conventional) but that representation is associated with the representations for their component parts.

The case we will focus on here is the creation of constructions that carry specific pragmatic implications while still maintaining at least a surface resemblance to the construction from which they arose. The example is a construction discussed in Fillmore and Kay 1994 and in C. Johnson 1997. Fillmore and Kay provide the following dialogue from a popular joke.

- (13) Diner: Waiter, what's this fly doing in my soup?
 Waiter: Why, madam, I believe that's the backstroke.
 (From Fillmore and Kay 1994)

The joke shows the possible ambiguity of the highlighted sequence. As Fillmore and Kay point out, the usual interpretation of 'what is X doing Y' is one of surprise at incongruity accompanied by more than a hint of disapproval. Because it is syntactically indistinct from the construction from which it arose, it gives the clever waiter license to interpret it as a literal question about what the fly is doing.

Since there is nothing in the form or content to suggest a meaning of incongruity, how did an ordinary Wh-question with doing and a locative phrase acquire these implications? C. Johnson (1997, forthcoming) points out that certain locative expressions have an ambiguity such that they can either apply to an activity or an individual participant in that activity (p. 17).

- (14) What are you doing with that knife = 'why do you have that knife?'
 or the literal meaning = 'what are you doing with it?'

In many contexts either interpretation is acceptable.

The implication of disapproval (a subjective interpretation) must have come from multiple instances of use with this negative nuance. As Johnson shows these implications have been around for a long time with little apparent change. A search of the quotations in the OED shows a construction with do in the simple present with these implications beginning in the 14th century. The example in 15 shows that the question is more about location than about doing.

- (15) c1430 Syr Tryam. 431 What do ye here, madam? Fro whens come ye?

The example in 16 shows lingering ambiguity of the construction.

- (16) 1470-85 MALORY Arthur XIV. v,
What dost thow here? He ansuerd I doo neyther good nor grete ylle.
 'What are you doing here? He answered I do neither good nor great ill.'

Clear implications of disapproval are evident later on:

(17) 1656 COWLEY Misc., Swallow 3

Foolish Prater, what do'st thou
 So early at my window do
 With thy tuneless Serenade?

Despite its close relation to regular Wh-questions, the construction has developed special implications that have been passed down for centuries. Let us examine the question of how these implications come to be attached to this particular string of words. (For an analysis of how children acquire the implications that go with this construction, see C. Johnson 1997).

The implication of disapproval becomes conventionalized with a Wh-question with do and a locative phrase because of the ambiguity mentioned earlier and the fact that the construction with do happens to occur often in a context where the question implies disapproval. It appears that listeners and speakers follow and keep track of the implications that occur in particular contexts. In order to know that a certain implication has occurred frequently and is associated with a certain string of words, speakers must register the context and the implications from the very first exposure. They could not wait until they had heard the expression frequently in a certain context to register this in memory, because if they did not remember each time, they would not know that they had heard it before. Thus it seems clear that in order for this construction to acquire its special meaning, the learner/hearer must record in memory the implication of incongruity and disapproval from the very first exposure. Because this construction was earlier just a special instance of a more general construction, the changes it has undergone indicate that particular instances of constructions are registered in linguistic memory indexed with their implications and contexts of use.

A further note of some interest is that in the 18th century, the construction was reworked with the progressive.

(18) 1835 Aristophanes' Clouds in Blackw. Mag. XXXVIII. 520

Str. Pray who's that in the basket hung up in the air?.. Do tell me, I
 pray, what you're doing up there.
Soc. Aerobating--sun-musing, pacing air.

This reformation with the progressive suggests that the construction was not frequent enough nor autonomous enough to maintain the old form with simple do. As Johnson has argued, this special construction maintains its relation with the more general Wh-question. Compare this to the high frequency greeting How do you do? which has retained the simple present form. In addition, it can be noted that the progressive, as a more recent grammaticization, maintains a fairly strong sense of action in progress and has other uses with negative affect (e.g. Why are you standing there? Why are you bothering me?) (Murray 1989).⁵ The reformation affirms that particular instances of general constructions can have their own representation while still being associated with the general construction.

12. Special phonological reduction.

Just as phrases or instances of constructions can have special implications, and special meaning, particular instances of constructions can also develop special phonology through frequency of use (Bybee and Scheibman 1999). Frequent phrases such as I don't know, I don't think, and many others show phonological reduction in excess of that which could be attributed to online reduction processes, indicating that such reduction has accumulated in representation. High frequency words and phrases have larger exemplar clusters and greater ranges of variation than low frequency words and phrases. As phonological reduction occurs on line, high frequency words and phrases have more opportunity to undergo reduction. When an already reduced exemplar is selected for production, it may undergo further reduction, leading to more advanced reduction in high frequency phrases (Bybee 2000b, 2001, 2002, Pierrehumbert 2001). Note that in order for special phonological reduction to build up for frequent phrases the whole phrase has to have a representation.

The traditional proposal for what is stored in memory is that only forms with idiosyncratic phonology, semantics, or morpho-syntax would be stored. From a diachronic point of view, however, it is clear that in order for a form to acquire special phonology, semantics, or morpho-syntax, it must already be stored in memory. Thus exemplar representation is necessary to account for the emergence of new constructions, the development of special phonology or pragmatic implications in particular instances of constructions, and the existence of familiar prefabs.

Sections 14 and 15 discuss two cases that illustrate the need for particular exemplars of constructions to accrue strength in memory representation through repetition or frequency. Before turning to those cases, however, an examination of the way an exemplar model can account for the phonological and semantic changes discussed in the preceding sections is necessary. The following section contains an expansion of exemplar models that allows exemplar reorganization to take place on both the form and meaning sides of a linguistic sign.

13. Exemplar representation of the linguistic sign.

Illustrations of exemplars models on linguistic phenomena to date have been one dimensional, focusing only on the phonetic representation of words and phrases. In the examples of grammaticization and the development of new constructions discussed in sections 11 and 12, however, we have also seen that exemplar representations are necessary for the meaning and pragmatic context of an utterance. Indeed, a linguistic sign (either a word or a construction) has both a form and a meaning. In addition, it requires an association between the form and the meaning. Using these examples of linguistic change, let us consider the nature of an exemplar representation of a linguistic sign.

We know that specific phonetic representations are associated in representation with specific meanings and contexts because in grammaticization phonological reduction only occurs in the grammaticizing construction. Thus the form gonna is associated with the intention/future meaning and not with the movement in space or purpose meaning. In Figure 2, circles represent exemplars and lines the association between form and meaning.

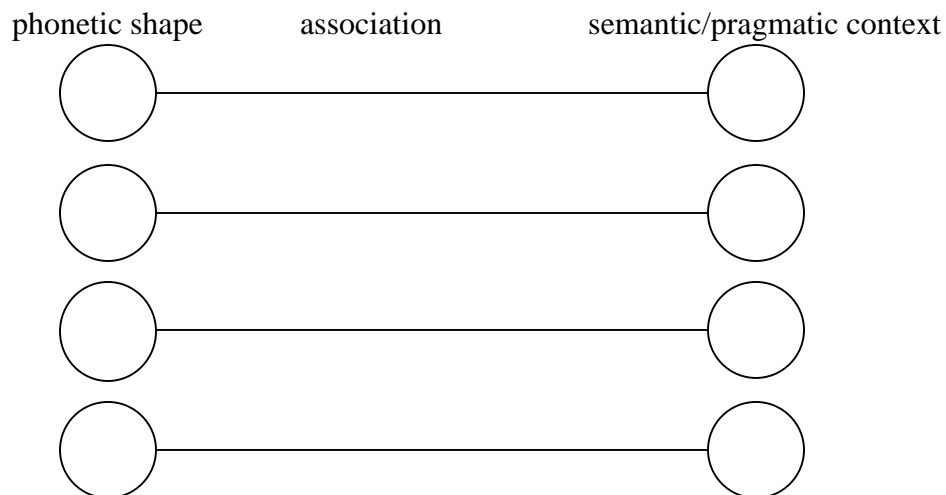


Figure 2: Four exemplars with phonetic shape associated uniquely with meaning and context.

Of course, in the variable phase of change, the situation will not be so neat. Some meanings will be associated with more than one phonetic shape and some phonetic shapes with more than one meaning as shown in Figure 3.

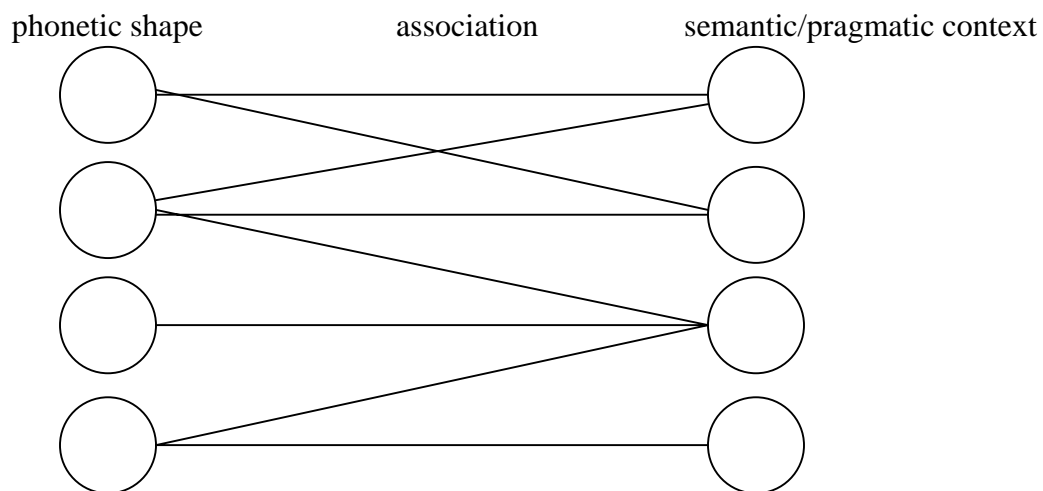


Figure 3: Variable associations of form and meaning in a linguistic sign

We have already noted that a cluster of phonetic exemplars may undergo reorganization with the more frequent exemplar taking over in cases where it was not originally conditioned by the phonetic environment (Bybee 2000a, Pierrehumbert 2002, Wedel 2004). The same can occur with the semantic/pragmatic exemplars. If a particular pragmatic inference, such as intention in the be going to example, occurs very frequently in the semantic cluster of exemplars with a particular form, that meaning can become firmly associated with the form and cause it to show up in new contexts, such as those in which intention, but no movement in space, is signified.

Similarly, the associations between form and meaning may also reorganize, with the most frequent associations becoming fixed and conventionalized. For instance, in the study of the reduction of don't, we found that forms with a full vowel in I don't know had a greater chance of occurring in a context where the meaning was literal or transparent, rather than contexts in which the phrase was used in a pragmatic sense as a conversational hedge or turn organizer (Bybee and Scheibman 1999; Scheibman 2000). However, the correspondence was not perfect: both the reduced and nonreduced vowel occurred in the transparent and pragmatic uses, as shown in Table 1.⁶ Given the strong tendency of the full vowel to occur with the transparent meaning, one could easily imagine a point in the future in which the two uses are distinguished phonologically.

	<u>Full vowel</u>	<u>reduced vowel</u>
transparent use	7	12
pragmatic use	1	17

Table 1. The relation between the transparent versus pragmatic uses of I don't know and the full versus reduced vowel variants.

Using the type of diagram shown in Figures 2 and 3, the associations between the reduced form of don't in the phrase I don't know and the transparent vs. pragmatic uses would appear as in Figure 4. The size of the circles corresponds to the frequency of the exemplar. The thickness of lines indicates the strength of the association, based on the numbers in Table 1.

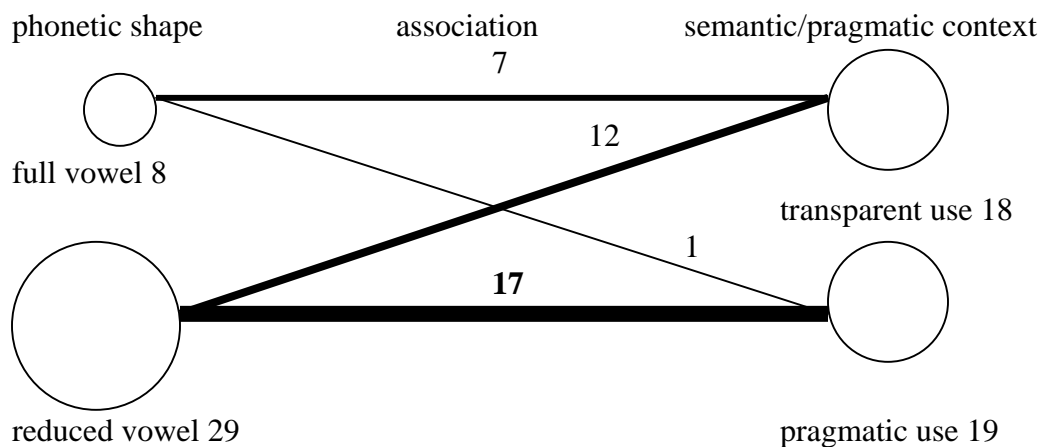


Figure 4: Phonetic and semantic associations in exemplars of I don't know (data from Bybee and Scheibman 1999).

Given the fact that infrequent exemplars tend to be lost, we might predict that the use of the full form in the pragmatic sense, which is already very unlikely, would eventually become nonexistent. In this way variable situations can be resolved towards a

situation in which an innovative phonetic form has a specific sense. Of course, this is precisely what occurs in the process of grammaticization.

Another way of conceiving of the form-meaning relation for exemplars would be to have each exemplar carry all the properties it had in use. Thus rather than positing a set of phonetic exemplars and a set of semantic/pragmatic ones with associations between them, each exemplar would be specified for both form and meaning. Then exemplars would overlap one another; if they have the same phonetic shape that part of the exemplar would overlap; if it were the meaning that was shared, then that part of the exemplar would overlap. Rare exemplars (such as the one in which don't is not reduced in I don't know but it has a pragmatic use) would eventually disappear. At present, I see no reason to choose one proposal over the other, except that I find the option chosen here easier to illustrate in two-dimensional space.

14. Frequency effects: category formation.

Continuing now with evidence that particular instances of constructions are stored in memory representations, we consider in this section and the next, two other effects of frequency of particular exemplars of constructions. The first case concerns that nature of the category formed by the lexical items that are used in the open slot in constructions. Sometimes this slot is highly generalized, taking all nouns or all verbs of the language. At other times, this slot is semantically constrained. In a study of the use of adjectives after certain verbs of becoming in Spanish, Bybee and Eddington found evidence that high frequency lexical instances of constructions can act as the central members of the category formed by members of the open slot in a construction. For two of the four verbs used in the study, the adjectives that were found with the particular verbs form an exemplar category with the most frequent member as the central member.

The study involved four Spanish verbs of 'becoming' used with human subjects and adjectives:

(19)	<u>quedarse</u> + ADJECTIVE	'to become (lit. to remain + refl.)'
	<u>ponerse</u> + ADJECTIVE	'to become (lit. to put oneself)'
	<u>hacerse</u> + ADJECTIVE	'to become (lit. to make oneself)'
	<u>volverse</u> + ADJECTIVE	'to become (lit. to turn oneself)'

The question addressed in many studies of these verbs in Spanish is how to determine which verb to use with a given adjective. Previous analyses have attempted to find general features that characterize the adjectives used with each verb: for example, is the change fast or slow, is the subject passive or active, is the result permanent or temporary? (Crespo 1949, Fente 1970)

Eddington 1999 studied these verbs in a large corpus that included both spoken and written materials. His results show that the features proposed by other researchers fail to distinguish the corpus examples of quedarse, ponerse, and volverse, all of which tend to be used with changes that are fast, passive, and temporary.

Looking again at the distribution of verb + adjective combinations in the corpus, Bybee and Eddington find that certain verb + adjective combinations have a high token frequency, for instance, the following, with their token frequencies per approximately 2 million words.

- | | | | |
|------|-------------------------|------|-------------------|
| (20) | <u>ponerse nervioso</u> | (17) | ‘to get nervous’ |
| | <u>quedarse solo</u> | (28) | ‘to end up alone’ |
| | <u>quedarse inmóvil</u> | (17) | ‘to be immobile’ |

For quedarse and ponerse, many other collocations were very similar to these semantically. For instance, quedarse inmóvil occurred 17 times in the corpus material (mostly in the written material) and its use as a prefab seems to have spawned many similar uses with other adjectives. In the corpus, there were fifteen other adjectives with meanings similar to inmóvil that were used only one to three times. These adjectives form a category with proto-type effects: the most frequent member is central and the other members are more marginal. Other members included adjectives that are basically synonymous with inmóvil such as parado ‘stopped, standing’; adjectives that are figurative/metaphorical de piedra ‘of stone’; and adjectives that share the ‘motionless’ feature but add other features, such as atrapado ‘trapped. Thus we argue for a category with the high frequency member as the center of the category.

Rather than attempting to find abstract semantic features that characterize all the adjectives used with a particular verb, Bybee and Eddington propose that the verbs quedarse and ponerse each have a number of such clusters of adjectives. For instance, central members of categories for quedarse besides the two mentioned above are the following, which are related to quedarse inmóvil-- quedarse quieto ‘to become still’, quedarse tranquilo ‘to calm down’, quedarse callado ‘to quiet down, become silent’, quedarse dormido ‘to fall asleep’ and two other categories that are unrelated, quedarse sorprendido ‘to be surprised’ and quedarse embarazada ‘to get pregnant’. See Bybee and Eddington for further details.

Three types of evidence suggest that the most frequent member is central to the category. First, the larger, more productive categories (that is, those with the highest type frequency) are organized around a frequent member, suggesting that productive uses of the verb occur on the basis of reference to semantic similarity to a frequent member. The groups of adjectives that occurred with the verbs that were less productive did not have a high frequency member. Of the four verbs studied, only two, quedarse and ponerse, showed categories organized around high frequency exemplars. The other two verbs, which are much less frequent with animate subjects, had a much more scattered and miscellaneous distribution with adjectives.

Second, family resemblance structure uses the most frequent adjective as central. For example, the chain in (21) takes inmóvil with its meaning of motionlessness as the link between the other two adjectives.

- | | | | | | |
|------|--------------------|---|----------------|---|-----------------|
| (21) | <u>petrificado</u> | - | <u>inmóvil</u> | - | <u>atrapado</u> |
| | ‘turned to stone’ | | ‘immobile’ | | ‘trapped’ |

Third, we conducted an experiment on the acceptability of verb + adjective combinations. We asked 48 peninsular Spanish speakers to rate the acceptability of sentences we took from the corpus. All stimuli were all naturally-occurring utterances of Spanish. The subjects were asked to rate the stimuli on a 5 point scale from ‘perfectamente bien’ ‘perfectly fine’ to ‘raro’ or ‘strange/rare’.

There were three groups of stimuli:

1. High frequency phrases (quedarse inmóvil)
2. Low frequency phrases with a close semantic affinity to a high frequency phrase (quedarse parado)
3. Low frequency phrases with no semantic connection to a high frequency phrase (quedarse orgullosísimo ‘become very proud’)

The results for quedarse are shown in Figure 5.

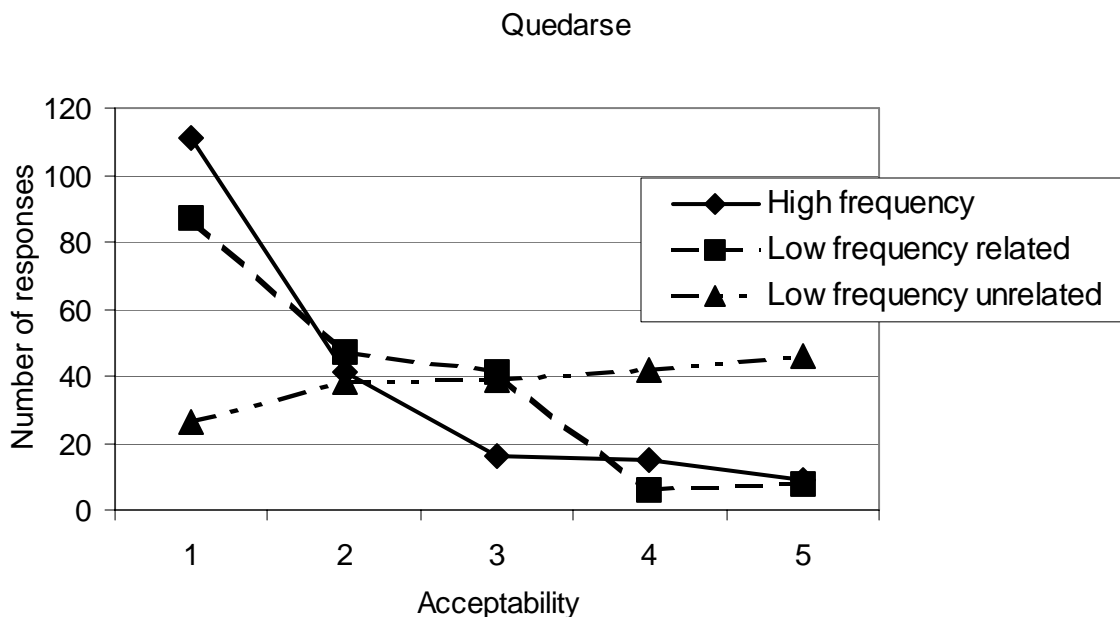


Figure 5. Acceptability judgments. ‘1’ is high acceptability and ‘5’ is low.

1. The high frequency phrases were judged most acceptable. ($X^2(1) = 51.4, p < .0001$)
2. A significant difference was found between the high frequency phrases and the low frequency phrases that were semantically similar to the high frequency phrases. (For quedarse and ponerse together: $X^2(1) = 6.22, p < .013$.)
3. The strongest result was the significant difference between the low frequency related items and the low frequency unrelated items. ($X^2(1) = 32.9, p < 0.0001$)

Thus both frequency and semantic similarity to a frequent exemplar significantly influenced the subjects’ judgments of acceptability. We can conclude that not only do speakers record specific exemplars of constructions in memory, but that frequency of use also has an impact on the strength of these exemplars and on category formation. Acceptability judgments rely on familiarity: frequent word sequences and word sequences similar to frequent ones will be judged more acceptable than low frequency, isolated combinations.

15. Frequency effects: resistance to change.

Exemplars of morpho-syntactic constructions, like morphologically complex words, are resistant to change if they are highly frequent. Just as irregular verbs that are of high frequency resist regularization, so constructions with odd or irregular properties resist change in the particular exemplars that are of high frequency. (See Ogura 1993, Bybee and Thompson 1997, Smith 2001).

Tottie 1991 studied the use of not-negation and no-negation (also called neg-incorporation) as in the examples such as the following.

- (22) He did not see anything. (not-negation)
 (23) He saw nothing. (no-negation)

Diachronically, the no-negation construction predates the construction with not and the more recently developed construction with not is more productive, gradually increasing its usage. Thus we can predict that the no-negation construction would be maintained primarily in high frequency constructions and collocations, while not-negation would be spreading to more contexts.

Tottie studied these two constructions in a large corpus of spoken and written British English. She extracted only those examples where the use of the alternate construction would have the same meaning and implications (as in 22 and 23). She found that certain constructions, especially existential be (as in example 24), stative have (as in 25) and copular be (as in 26) have a higher use of no-negation than lexical verbs do, as shown in Table 2. This suggests that no-negation, rather than being an option for all sentences, has become associated with certain constructions.

	Spoken		Written	
Existential <u>be</u>	34/38	89%	96/98	98%
Stative <u>have</u>	18/28	64%	41/42	98%
Copular <u>be</u>	12/20	60%	26/47	55%
Lexical verbs	20/76	26%	67/104	64%

Table 2. Proportion of no-negation (Tottie 1991)

- (24) by the time they got to summer there was no more work to do
 (25) the Fellowship had no funds
 (26) as a nation we are not doing well enough. This is no new discovery

These three constructions are fairly frequent, accounting together for more of the data than all the lexical verbs combined. Their frequency could help explain the fact that they preserve the older construction; much like the old ablauting verbs of English (break, broke; write, wrote; etc.) their high frequency strengthens their representations and makes them less likely to be reformed on the more productive pattern. This suggests that a frequency effect might also be found among the lexical verbs.

In fact, certain frequent verbs, i.e. know, do, give, make account for many of the examples of no-negation in the lexical examples:

- (27) no, Marilyn does no teaching I imagine she's a research assistant
 (28) I've done nothing except you know bring up a family since I left school
 (29) I know nothing about his first wife

In addition, some lexical verbs occur in conventionalized expressions or prefabs that are mostly used in writing:

- (30) the ballads make no mention of the trapping of rabbits
 (31) Make no mistake about it, the divisions are very serious
 (32) the split in the Conservative Party over Africa gives me no joy

These examples demonstrate that even after a construction has lost its productivity, specific exemplars of the construction may live on because they have accrued strength through repetition and so continue to be used.

16. Conclusion. The preceding discussion has presented a number of cases which argue, each one in a different way, for memory or cognitive storage of particular exemplars of constructions. This evidence includes the following facts:

1. Speakers are familiar with certain word combinations (prefabs) which are in no way exceptional in meaning or form.
2. Special phonological reduction accrues to specific high frequency phrases.
3. New constructions are created out of specific instances of old general constructions.
4. In grammaticization, changes in phonology, semantics, and structure occur in extremely high frequency constructions.
5. Certain higher frequency exemplars of constructions dominate formation of categories of items within constructions.
6. High frequency exemplars of constructions resist change on the basis of more productive constructions.

These facts all point to the conclusion that grammar is the cognitive organization of experience with language. Usage feeds into the creation of grammar just as much as grammar determines the shape of usage. Actual language use cannot be omitted from the study of grammar, because it constitutes a large part of the explanation for why languages have grammar and what form that grammar takes.

Grammar cannot be thought of as pure abstract structure that underlies language use: just as there can be no discrete separation of grammar and lexicon because there are so many cases in which specific lexical items go with / require certain grammatical structures, there can be no strict separation of grammar and usage. Grammar is built up from specific instances of use which marry lexical items with constructions; it is routinized and entrenched by repetition and schematized by the categorization of exemplars.

A conceptualization of grammar as pure structure fails to provide us with explanations for the nature of grammar. A theory based on usage, on the other hand, which takes grammar to be the cognitive organization of language experience can reference general cognitive abilities: the importance of repetition in the entrenchment of

neuro-motor patterns, the use of similarity in categorization and the construction of generalizations across similar patterns. These processes, combined with the functions of language in context, such as establishing reference, maintaining coherence, signaling turn-taking, explain grammar as the ritualization of oft-repeated routines.

Language can be viewed as a complex system in which the processes that occur in individual usage events, such as those just mentioned, with high levels of repetition, not only lead to the establishment of a system within the individual, but also lead to the creation of grammar, its change, and its maintenance within a society.

References

- Anderson, John R. 1993. *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.
- Barlow, Michael. 2000. Usage, blends and grammar. In *Usage-based models of language*, ed. by Michael Barlow and Suzanne Kemmer, 315-45. Stanford: CSLI.
- Barlow, Michael, and Suzanne Kemmer (eds.). 2000. *Usage-based models of language*. Stanford: CSLI.
- Biber, Douglas, S. Johansson, G. Leech, S. Conrad, E. Finegan. 1999. *Longman grammar of spoken and written English*. Harlow, England: Longman.
- Bod, Rens. 1998. *Beyond grammar: an experience-based theory of language*. Stanford: CSLI.
- Bolinger, Dwight. 1961. Syntactic blends and other matters. *Language* 37.366-81.
- Boyland, Joyce Tang. 1996. *Morphosyntactic change in progress: a psycholinguistic approach*. Berkeley, CA: University of California dissertation.
- Bybee, Joan. 1985. *Morphology: a study of the relation between meaning and form*. Amsterdam, Philadelphia: John Benjamins.
- Bybee, Joan. 1998. The emergent lexicon. *Chicago Linguistic Society* 34.421-35.
- Bybee, Joan. 2000a. Lexicalization of sound change and alternating environments. *Papers in laboratory phonology V: acquisition and the lexicon*, ed. by Michael Broe and Janet Pierrehumbert, 250-68. Cambridge: Cambridge University Press.
- Bybee, Joan. 2000b. The phonology of the lexicon: evidence from lexical diffusion. *Usage-based models of language*, ed. by Michael Barlow and Suzanne Kemmer, 65-85. Stanford: CSLI.
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, Joan. 2003a. Mechanisms of change in grammaticization: the role of frequency. *The handbook of historical linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 602-23. Oxford: Blackwell.
- Bybee, Joan. 2003b. Cognitive processes in grammaticalization. *The new psychology of language, volume II*, ed. by Michael Tomasello. Mahwah, NJ: Lawrence Erlbaum.
- Bybee, Joan, and David Eddington. (in preparation). A usage-based exemplar model approach to Spanish verbs of 'becoming'.
- 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 degrees of constituency: the reduction of *don't* in English. *Linguistics* 37.575-596.
- Bybee, Joan, and Sandra A. Thompson. 1997. Three frequency effects in syntax. *Berkeley Linguistics Society* 23.378-88.
- Chandler, Steve. 2002. Skousen's analogical approach as an exemplar-based model of categorization. *Analogical modeling*, ed. by Royal Skousen, Deryle Lonsdale, and Dilworth B. Parkinson, 51-105. Amsterdam: John Benjamins.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Crespo, Luis A. 1949. To become. *Hispania* 32.210-12.
- Croft, William. 2001. *Radical Construction Grammar: syntactic theory in typological perspective*. Oxford University Press.
- Danchev, Andrei, and Merja Kyto. 1994. The Construction *be going to* + Infinitive in Early Modern English. *Studies in Early Modern English*, ed. by Dieter Kastovsky, 59-77. Berlin: Mouton de Gruyter.
- Eddington, David. 1999. On 'becoming' in Spanish: a corpus analysis of verbs of expressing change of state. *Southwest Journal of Linguistics* 18.23-46.

- Erman, Britt, and Beatrice Warren. 2000. The Idiom Principle and the Open Choice Principle. *Text* 20.29-62.
- Fente, R. 1970. Sobre los verbos de cambio o devenir. *Filología Moderna* 38.157-72.
- Fidelholtz, James L. 1975. Word frequency and vowel reduction in English. *Chicago Linguistic Society* 11.200-13.
- Fillmore, Charles, and Paul Kay. 1994. Grammatical constructions and linguistic generalizations: the What's X doing Y? construction. Ms.
- Fillmore, Charles J., Paul Kay, and Mary C. O'Connor. 1988. Regularity and idiomaticity in grammatical constructions. *Language* 64.501-38.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Goldberg, Adele. 1995. *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, Adele E. 2003. Constructions: a new theoretical approach to language. *Trends in Cognitive Science* 7.219-24.
- Goldinger, Stephen D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition* 22.1166-83.
- Greenberg, Joseph. 1966. *Language universals: with special reference to feature hierarchies*. The Hague: Mouton.
- Gregory, Michelle L., William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. *Chicago Linguistic Society* 35.151-66.
- Haiman, John. 1994. Ritualization and the development of language. *Perspectives on grammaticalization*, ed. by William Pagliuca, 3-28. Amsterdam, Philadelphia: John Benjamins.
- Haspelmath, Martin. 1998. Does grammaticalization need reanalysis? *Studies in Language* 22.315-51.
- Hay, Jennifer. 2001. Lexical frequency in morphology: is everything relative? *Linguistics* 39.1041-70.
- Heine, Bernd, Ulrike Claudi, and Frederike Hunnemeyer. 1991. *Grammaticalization: a conceptual framework*. Chicago: University of Chicago Press.
- Hooper, Joan. 1976. Word frequency in lexical diffusion and the source of morphophonological change. *Current progress in historical linguistics*, ed. by William Christie, 96-105. Amsterdam: North-Holland.
- Hopper, Paul. 1987. Emergent grammar. *Berkeley Linguistic Society* 13.139-57.
- Hopper, Paul, and Sandra A. Thompson. 1980. Transitivity in grammar and discourse: *Language* 56.251-99.
- Hopper, Paul, and Elizabeth Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Israel, Michael. 1996. The way constructions grow. *Conceptual structure, discourse and language*, ed. by Adele E. Goldberg, 217-30. Stanford: CSLI Publications.
- Johnson, Christopher. 1997. The acquisition of the 'What's X doing Y?' construction. *Proceedings of the 21st annual Boston University Conference on Language Development* 2.343-53. Somerville, MA: Cascadilla Press.
- Johnson, Christopher. Forthcoming. The functions of ambiguity: towards a developmental theory of meaning extension.
- Johnson, Keith. 1997. Speech perception without speaker normalization. *Talker variability in speech processing*, ed. by Keith Johnson and John W. Mullennix, 145-65. San Diego: Academic Press.
- Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. Probabilistic relations between words: evidence from reduction in lexical production. *Frequency and the emergence of linguistic structure*, ed. by Joan Bybee and Paul Hopper, 229-54. Amsterdam: John Benjamins.
- Kay, Paul, Charles Fillmore, Laura A. Michaelis, and Ivan A. Sag. Forthcoming. *Construction grammar*. Stanford: CSLI.
- Lakoff, George. 1987. *Women, fire and dangerous things*. Chicago: University of Chicago Press.
- Langacker, Ronald. 1987. *Foundations of cognitive grammar, Vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald. 2000. A dynamic usage-based model. *Usage-based models of language*, ed. by Michael Barlow and Suzanne Kemmer, 1-63. Stanford: CSLI.
- Larsen-Freeman, Diane. 1997. Chaos/complexity science and second language acquisition. *Applied Linguistics* 18.141-65.

- Lieven, Elena, Michael Tomasello, Heike Behrens, and Jennifer Speares. 2003. Early syntactic creativity: a usage-based approach. *Journal of Child Language* 30.333-70.
- Lindblom, Björn, Peter MacNeilage, and Michael Studdert-Kennedy. 1984. Self-organizing processes and the explanation of phonological universals. *Explanations for language universals*, ed. by Brian Butterworth, Bernard Comrie, and Östen Dahl, 181-203. New York: Mouton.
- Murray, Denise E. 1989. Social and functional uses of the present progressive. *TESL Reporter* 22.72-77.
- Newmeyer, Frederick J. 2003. Grammar is grammar and usage is usage. *Language* 79. 682-707.
- Nosofsky, Robert M. 1988. Similarity, frequency and category representation. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14.54-65.
- Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow. 1994. Idioms. *Language* 70.491-538.
- Ogura, Mieko. 1993. The development of periphrastic do in English: a case of lexical diffusion in syntax. *Diachronica* X.51-85.
- Ono Tsuyoshi, Ryoko Suzuki, and Sandra A. Thompson. 2000. The pragmatic nature of the so-called subject marker ga in Japanese: evidence from conversation. *Discourse Studies* 2.55-84.
- Pawley, Andrew, and Frances Hodgetts Syder. 1983. Two puzzles for linguistic theory: nativelike selection and nativelike fluency. *Language and communication*, ed. by Jack C. Richards and Richard W. Schmidt, 191-226. Longman: London.
- Pierrehumbert, Janet. 2001. Exemplar dynamics: word frequency, lenition and contrast. *Frequency and the emergence of linguistic structure*, ed. by Joan Bybee and Paul Hopper, 137-58. Amsterdam, Philadelphia: John Benjamins.
- Pierrehumbert, Janet. 2002. Word-specific phonetics. *Laboratory phonology 7*, ed. by C. Gussenhoven and N. Warner, 101-39. Berlin: Mouton de Gruyter.
- Poplack, Shana. 2001. Variability, frequency and productivity in the irrealis domain of French. *Frequency effects and emergent grammar*, ed. by Joan Bybee and Paul Hopper, 405-28. Amsterdam: Benjamins.
- Poplack, Shana, and Sali Tagliamonte. 1999. The grammaticization of going to in (African American) English. *Language Variation and Change* 11.315-42.
- Poplack, Shana, and Sali Tagliamonte. 2001. *African American English in the diaspora*. Oxford: Blackwell.
- Savage, Ceri, Elena Lieven, Anne Theakston, and Michael Tomasello. 2003. Testing the abstractness of young children's linguistic representations: Lexical and structural priming of syntactic constructions? *Developmental Science* 6.557-67.
- Scheibman, Joanne. 2000. I dunno but... a usage-based account of the phonological reduction of don't in conversation. *Journal of Pragmatics* 32.105-24.
- Schuchardt, H. 1885. *Über die Lautgesetze: Gegen die Junggrammatiker*. Berlin: Robert Oppenheim.
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer Academic.
- Smith, K. Aaron. 2001. The role of frequency in the specialization of the English anterior. *Frequency effects and emergent grammar*, ed. by Joan Bybee and Paul Hopper, 361-82. Amsterdam, Philadelphia: John Benjamins.
- Thompson, Sandra A., and Paul J. Hopper. 2001. Transitivity, clause structure, and argument structure: evidence from conversation, *Frequency effects and emergent grammar*, ed. by Joan Bybee and Paul Hopper, 27-60. Amsterdam, Philadelphia: John Benjamins
- Tomasello, Michael. 2003. *Constructing a language*. Cambridge, MA: Harvard University Press.
- Tottie, Gunnel. 1991. Lexical diffusion in syntactic change: frequency as a determinant of linguistic conservatism in the development of negation in English. *Historical English syntax*, ed. by Dieter Kastovsky, 439-67. Berlin: Mouton de Gruyter.
- Traugott, Elizabeth C. 2003. Constructions in grammaticalization. *The handbook of historical linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 624-47. Oxford: Blackwell.
- Wedel, Andrew. 2004. Category competition drives contrast maintenance within an exemplar-based production/perception loop. *Proceedings of the Workshop of the ACL Special Interest Group in Computational Phonology*.
- Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

¹ A large corpus provides only an approximation of a speaker's experience with language, which of course will be different for different speakers.

² See Barlow 2000 for an interesting discussion of the way a conventionalized expression can undergo permutations that demonstrate that its compositionality is also maintained.

³ In some versions of exemplar representation exemplars are scattered randomly through space. Only when categorization of a new exemplar is necessary are they organized by similarity (Chandler 2002, Skousen 1989). Because linguistic categorization takes place so often I propose that linguistic categories (in the form of phonetic, morphosyntactic, semantic/pragmatic characteristics) are more entrenched in the sense that frequently used categorizations have an impact on neurological organization.

⁴ Shakespeare's works also contain a number of purpose clauses with *-ing* forms without a copula functioning as subordinate clauses:

Besides, I met Lord Bigot and Lord Salisbury,
 With eyes as red as new-enkindled fire,
 And others more, going to seek the grave
 Of Arthur, who they say is kill'd to-night
 On your suggestion. (King John, IV.1)

⁵ When this paper was delivered as the Presidential Address in 2005, Ray Jackendoff asked why this particular construction reformed with the progressive, while other formulaic constructions maintain older morphological forms, such as the phrase far be it from me. He asked why one reforms and the other does not. This is partly explained as noted in the text by the fact that the progressive, as a newly grammaticized form with its 'in progress' semantics, is particularly appropriate for the WXYD construction, while changing the subjunctive to indicative in the far be it from me case has no particular advantage. Also, far be it from me is a discourse marker that does not have open slots for other lexical items and is primarily if not exclusively used with the first person singular. That is, it is much more fixed than the WXYD construction.

⁶ See Scheibman 2000 for further details. The 'turn organizer' use seems to be only associated with the reduced form.