CHAPTER 4

USAGE-BASED THEORY AND EXEMPLAR REPRESENTATIONS OF CONSTRUCTIONS

JOAN L. BYBEE

4.1. Introduction: Usage-based Theory

The basic premise of Usage-based Theory is that experience with language creates and impacts the cognitive representations for language (Langacker 1987, 2000*b*; Kemmer and Barlow 2000). Cognitive representations are built up as language users encode utterances and categorize them on the basis of phonetic form, meaning, and context. As incoming utterances are sorted and matched by similarity to existing representations, units such as syllable, word, and construction emerge. Thus, grammar can be viewed as the cognitive organization of one's experience with language (Bybee 2006).

It is accepted within this approach that it is wiser to begin the search for the cognitive processes involved in language by first considering domain-general cognitive processes—that is, those known to function in domains other than language, such as vision or neuromotor processing—rather than to assume a priori that language requires special adaptations of cognitive functions (Elman and Bates 1997; Tomasello 2003; Bybee and Beckner 2009). Some of these domain-general processes are categorization, cross-modal association,

and neuromotor automation. Consider how these processes apply to language. Categorization applies at every level of form and meaning: phones and their combinations are categorized based on existing representations, as are features of context and meaning (Langacker 2000*b*; Bybee 2010). Cross-modal association allows humans to match up the phonetic (or manual) form experienced with properties of the context and meaning. Automation allows the units of language to be combined in conventional ways that lead to fluency in both production and perception (Bybee 2002).

An important characteristic of human language is that the individual units and sequences of units are subject to high levels of repetition. It is repetition that leads to conventionalization of categories and associations, as well as to the automation of sequences. Because some units and sequences are repeated more than others, it has been possible to identify the properties of cognitive representations that depend upon the extent to which they have been accessed for production or perception. Thus, within Usage-Based Theory the study of frequency effects of various sorts has contributed to the understanding of the nature of grammatical organization (Bybee 2007).

It is recognized that languages are constantly changing, and this change is gradual and takes place as language is used. Change is attributed to the way particular cognitive processes apply in language use, thus change provides an important window into the understanding of the cognitive processes underlying language. As change is gradual, the categories and units of language are variable and they form gradient rather than strictly bounded categories. Thus, linguistic structure is viewed as emergent—governed by certain regular processes, but always changing as it is re-created in the individual and in specific usage situations (Hopper 1987; Kemmer and Barlow 2000; Ellis and Larsen-Freeman 2006*b*). Thus, rather than a fixed, static set of representations, language is viewed as being affected by experience in an ongoing way even in adults. It also follows that we should not expect linguistic constructs such as segment, syllable, morpheme, word, or construction to have strict definitions, nor do we expect all the manifestations of these constructs in languages to exhibit exactly the same behavior (Bybee 2010: chapter 1).

As the term implies, the object of study in Usage-based Theory is not only the native user's competence, which resides in the cognitive representation, but all the perception and production processes that are brought to the task of using language (Kemmer and Barlow 2000). Thus, the data considered applicable to formulating and testing hypotheses can be very broadly drawn from experiments, child language acquisition, language change, and large corpora representing natural usage. The latter source of data (now much more available than in the past) partially supplants the use of native speaker intuitions. While such intuitions are interesting and important, research with natural usage indicates that language users are often unaware of the nature and frequency of certain structures that they use. For this reason, it is impor-

tant to ground linguistic research in recorded instances of language use in natural situations.

The term 'Usage-based' was coined by Langacker (1987), but the roots of this view of language can be found starting in the 1960s with the functional-typological approach to language taken by Joseph Greenberg and colleagues, and then flowering in the 1970s under the leadership of Givón, Li, Thompson, Hopper, and many others (see, e.g., Givón 1979; Li 1976; Hopper and Thompson 1980, 1984). These linguists proposed that grammar is created by the conventionalization of commonly used discourse patterns. This proposal, while not explicitly cognitive, encapsulates the usage-based premise that linguistic structure is formed by the repetition of certain linguistic patterns in language use. The proposal is enriched by the recognition of the cognitive processes that feed into linguistic structure, such as categorization, as mentioned above, but also entrenchment and schematization (Langacker 1987, 2000*b*), metaphor (Lakoff 1987), inference, and other types of construal (Traugott 1989).

4.2. Constructions and Usage-based Theory

Constructions, with their direct pairing of form to meaning without intermediate structures, are particularly appropriate for usage-based models. As Noonan (1998) observes, functionalist usage models are materialist, 'what-you-see-is-what-you-get' models, in which the language user's experience with language is represented rather directly in cognition. It would appear not to be entirely accidental that research on constructions and research on usage and cognitive effects have converged on this view in recent years.

From a grammarian's point of view, constructions are identifiable because they are groupings of words that have idiosyncratic behavior at some level: they might be formally special, but more often they take on an unpredictable meaning or pragmatic effect (Fillmore, Kay, and O'Connor 1988; Goldberg 1995). From the broader perspective of usage-based theory, however, constructions can be viewed as processing units or chunks—sequences of words (or morphemes) that have been used often enough to be accessed together. This would mean that word sequences that are often used are constructions even if they do not have idiosyncrasies of meaning or form (Bybee 2001: 173, 2006; Goldberg 2006*a*). It is interesting, of course, that such chunks or conventionalized sequences have a tendency over time to develop special pragmatic implications that can lead to special meaning. They can also develop idiosyncrasies of form in a variety of ways.

4.3. Exemplar Representation for Constructions

Despite the more surface-oriented view of Construction Grammar in comparison to generative grammar in its various manifestations over the decades, abstraction away from particular tokens does take place, as evidenced by innovative uses of constructions. One important question addressed in recent literature is the nature of this abstraction and how the experienced tokens of a construction contribute to the formation of a cognitive representation of that construction (Goldberg 2006*a*; Bybee 2010). This chapter discusses these questions taking an exemplar view of constructions.

The nature of exemplar representations will be explained in more detail in the next section, but for present purposes we can say that exemplar models propose that memory for linguistic experience is like memory for other types of experience: each token of experienced linguistic behavior has an impact on cognitive representation; when stored representations are accessed in either encoding or decoding, the representations themselves change. In addition, memory storage for linguistic experience includes detailed information about the tokens that have been processed, including their form and the contexts in which they were used. Exemplar representations contrast with the more abstract representations of structural or generative theories (at all levels—phonetic, morphosyntactic, and semantic/pragmatic), in that variation and features predictable from general principles have not been removed. In such a model, the general categories and units of grammar can emerge from the experience that is recorded in memory because exemplars are categorized by similarity to one another and because contiguous experiences—such as meaning and acoustic shape—are recorded as linked to one another.

The following sections will describe the general properties of exemplar models and discuss how they can be applied to constructions. The following arguments for choosing exemplar representations for constructions will be presented in this chapter:

1. As Fillmore et al. (1988) argue, many constructions have idiosyncratic features of morphosyntax, semantics, pragmatics, and phonology, and much of what a speaker/hearer knows about his/her language is not predictable from the very general rules that have occupied the attention of most syntacticians in the past, but rather consists of specific information that must be associated with specific constructions. Specific information finds a natural expression in an exemplar model, where the storage and categorization of all detail both predictable and idiosyncratic is considered to be a basic response to linguistic input and applies to all constructions, specific or general (see sections 4.4 and 4.5).
2. The stored representations in the form of exemplars respond to usage by allowing the representation of both token and type frequency; these

frequency patterns are important for understanding the categories that are formed for the schematic slots in constructions (see section 4.6).

3. Specific instances of constructions develop into new constructions, thus specific exemplars of constructions need to have cognitive representation (see section 4.7).
4. Exemplar models allow for specific meaning from the context of use to impact cognitive representation, which then accounts for the way that words and constructions are affected by the meaning that occurs in the context. The impact of context is seen both in the development of special meanings and implications of constructions and also in developments such as the negative prosody identified in corpora for constructions with *cause* (see section 4.8).
5. Exemplar models were not developed especially for language; rather they apply equally to linguistic and nonlinguistic categories. Such models then take the usage-based view that language is a part of general cognition and allow us to access explanations for linguistic phenomena outside of language.

4.4. Exemplars of Construction: Constructions as Chunks

Exemplars are categories formed from tokens of experience that are judged to be the same (Pierrehumbert 2001). Linguistic exemplars come in a variety of sizes, ranging from a single segment, such as a vowel, to whole paragraphs, such as the Pledge of Allegiance. The exemplars themselves are grouped together by similarity. Thus, the vowels of *hit, swim*, and *sip* may be grouped together, the different phonetic realizations of a word, such as *pretty* will be grouped together, as well as exemplars for longer sequences, such as *all of a sudden*. These exemplar clouds, as they are called, constitute categories. Exemplar categories are structured by similarity and frequency (Nosofsky 1988) and often exhibit prototype effects. Because they are grouped together based on similarity, in theory any type of category could be represented in exemplars. However, it is common to find prototype effects emerging from categories because of differences in degrees of similarity.

In addition, exemplars may differ in strength depending upon the number of tokens that comprise them. That is, exemplars built up from a large number of tokens will be represented more strongly than those built up from a smaller number of tokens. The stronger exemplar or set of exemplars often forms the center of a category and other exemplars are more or less similar to the stronger exemplar or set of exemplars (Pierrehumbert 2001).

As exemplars are based on perceptual stimuli, and exemplars are grouped together based on similarity, we distinguish exemplar categories formed by

different criteria: thus exemplar clouds can be formed on phonetic, semantic, pragmatic, or contextual criteria. For any word, phrase, or construction, exemplar categories from these different domains are linked. Thus, an exemplar model of language must allow for links across domains, based on co-occurrence in experience. Such links create the form-meaning correspondences that constitute constructions. Thus, exemplars, like constructions, provide for direct pairings of form with meaning without intermediate representations (such as phrase structure or phonemic representations).

In addition to cross-domain linking, mapping, or linking also occurs among units in different contexts. Just as a category may be formed over the vowels in *hit, swim*, and *dip,* a category may be formed at word level despite the fact that a word may occur in many different contexts. In fact, the exemplar cloud of a word would include all the meanings and contexts in which the word has been experienced.

A linguist's first reaction to such a model is that it includes way too much information and would require way too much storage capacity. There are two responses to this objection: first, neural capacity is much greater than previously believed, and second, information is stored in a highly structured and efficient way. Highly structured storage of information results when categories are formed and similar items are stored in proximity to one another. Also, as we know from experience, memories for all types of sensory input can be lost (through forgetting) when that information is not reinforced by repetition or recency. Of greatest interest to linguists, however, is the question of how much generalization and abstraction occurs in the organization of linguistic experience. This is an empirical question that the model allows us to pose and investigate in a realistic way, as we see in the following sections.

Memory storage of complex units such as idiomatic phrases or constructions requires links of various sorts. First, there is the sequential linking that comes about through repetition of sequences of units. In the domain-general process of chunking, repeated sequences of elements, be they linguistic or not, are represented together as units that can be accessed directly rather than formed compositionally (Newell 1990; Ellis 1996; Bybee 2002). By these means repeated sequences become more fluent. Within a chunk, sequential links are graded in strength based on the frequency of the chunk or perhaps the transitions between the elements of a chunk. A construction is a chunk even though it may contain schematic slots, that is, the elements of a chunk can be interrupted.

Second, there are links from the individual elements of a chunk to those elements in other contexts. Even though a phrase such as *all of a sudden*, or a construction such as *drive* SOMEONE *crazy, mad, up the wall…* forms a chunk, that does not mean that the items that compose it are not still analyzable as words that occur elsewhere in cognitive representation. Analyzability (Langacker 1987) can be represented as links from the exemplar cluster of a word within a construction to the general exemplar cluster of the word. An interesting phenomenon is the loss of analyzability of chunks, a process that occurs gradually. As chunks are more frequently used, and accessed directly, they become more autonomous and their

component words can lose their association with exemplars of the etymologically same word (Bybee 2003; Beckner and Bybee 2009; Bybee 2010). In grammaticalization this process is referred to as decategorialization (Hopper 1991). For instance, in the phrase *in spite of*, the erstwhile noun, *spite*, has lost its ability to function as a noun, as in this phrase it is not modifiable. It has also lost its earlier meaning of 'in defiance' and adds nothing to the general meaning of concessive that the phrase now has. This loss has taken place gradually, which means that there are degrees of linking between words in different contexts (Beckner and Bybee 2009). Note that complex prepositions such as *in spite of* are not extremely frequent. As Hoffmann (2004, 2005) has pointed out, loss of analyzability can come about without high levels of frequency, encouraged by changes in distribution, meaning, and function.

Another aspect of chunking that is important for the understanding of how constructions emerge is that both nonlinguistic and linguistic chunks tend to have meaning assigned to them at the highest level possible (Ellis 1996; Bybee 2010). For instance, one might witness the repetition of a sequence of events such as a human throwing a ball, a dog running after it, catching it, and returning to the human, setting the ball at the human's feet. The human picks up the ball and throws it again. It is a human propensity to try to understand such sequences as chunks of behavior and to assign labels, such as 'playing fetch' to such chunks. Similarly, sequences of linguistic units that occur together repeatedly tend to be assigned meanings as a whole rather than simply as a sum of the parts, as can be seen in many expressions such as *in spite of, here and there*, or *all of a sudden,* which no longer are fully compositional. It is this chunking and labeling phenomenon that provides language with constructions.

4.5. Details in Exemplar Representation

4.5.1 Details of Semantics and Pragmatics

The original motivation for proposing constructions as part of grammar, as expressed in Fillmore et al. (1988), was that constructions have idiosyncratic structure and meaning. As mentioned above, this attribute is taken by some to be criterial in defining constructions, though in the usage-based view construction can be defined simply as frequently used and thus conventional word sequences. As mentioned above, details of form and usage are automatically registered in exemplar representation and through entrenchment can become an inherent part of the construction. In this section, based on Bybee (2006), I demonstrate that an exemplar model is essential to explaining how constructions acquire idiosyncratic semantic/pragmatic and phonetic properties. I also argue that redundant detail may turn out to be important in forming the bases of new uses of constructions.

In discussing the *What's X doing Y?* (WXDY) construction, Fillmore and Kay (1999) note that one reason for regarding this as a construction is its special pragmatic implications, which in some cases occur without any accompanying literal meaning, as in (1).

(1) What's that box doing up there?

As for the origins of this sense for the construction, they say:

> While the WXDY construction may have had its origin in conversational implicatures—through situations in which an individual A is clearly up to no good and B asks what A is doing—the semantics of incongruity is now CONVENTIONALLY associated with the special morphosyntax of WXDY constructs. (Fillmore and Kay 1995: 5; emphasis original)

The conventionalization of implicature (or from the hearer's point of view, inference) is also well-known from grammaticalization research (Bybee 1988*b*; Traugott 1989; Bybee et al. 1994; Traugott and Dasher 2002). It is thought that the frequent co-occurrence of an inference with a particular construction can lead to that inference being taken as part of the meaning of the construction. The originally inferential meaning can even replace the earlier meaning.

If we consider how this conventionalization occurs, we see that we need an exemplar model to account for it. In a model in which semantic representations contain only a core or abstract meaning and inferences are calculated on the fly in each context, there is no way for an implication/inference to stick to a construction. However, an exemplar model would record the inferences made in each instance of use, and if the same inference is made on multiple occasions, the strength of that inference increases. With sufficient strength among the semantic exemplars for a construction, the inference can become conventionalized as part of the meaning of the construction. This can happen over diachronic time and it can also happen over the course of language acquisition, as the learner establishes what the implications of acquired constructions are. Note that the speaker/hearer cannot wait until s/he has experienced the inference more than once to store it in memory, because there is no way to know if the inference has been experienced before if it is not registered in memory from the very first experience (Bybee 2006).

4.5.2. Phonetic Detail

Similarly, phonetic change can accrue to the words of a construction through instances of use in reducing contexts. Grammaticalizing constructions undergo reduction as frequency of use increases. Words in grammaticalizing constructions often have a wide range of variation, as with the phonetic variants of English *going to*, which range from the full form to variants spelled *gonna* to extreme reduction as in *I'm gonna* [aimə̃nə] The entrenchment of the reduction in the construction helps to distinguish one construction from another, as in the case of *used to* as

found in the following two examples, where the first, with past habitual meaning, has the phonetic shape [justə] while the second has the form [juzdʊ]. (Examples from the Time Magazine Corpus, 2000.[1])

(2) it is becoming difficult to distinguish between what **used to** be considered elite culture and mass culture
(3) All that money would be **used to** battle not only the drug traffickers but also the guerrillas who are aligned with them

Bybee and Scheibman (1999) argue that phonetic reduction and assimilation take place in production and higher frequency words and phrases have more opportunities to undergo these processes. The reduction or assimilation that occurs in production is minimal, but in an exemplar model its effects are registered in memory through the establishment of a new exemplar. If that exemplar is chosen for production later on, and altered more in the process, yet another exemplar will be established (Bybee 2000; Pierrehumbert 2001). For very high frequency words that occur in low stress positions, this gradual process can lead to extreme degrees of reduction, such as those which occur with *I'm going to* and *I don't know.* As in the case of pragmatic/semantic change, an exemplar model is essential to explaining how these changes come about.

Some such changes are quite subtle and can only be detected by instrumental study over corpus tokens. For example, in a large corpus study, Berkenfield (2001) found phonetic differences in the vowel of *that* which included both duration and degree of centralization, according to whether it was used as a demonstrative pronoun, a demonstrative adjective, a complementizer, or a relative clause marker. In other words, *that* differed according to the construction it was used in. Similarly, Hay and Bresnan (2006) find that the raising of /æ/ in New Zealand English is more advanced in the word *hand* when it refers to a limb than when it is used in phrases such as *give a hand*. They also find that the centralization of /ɪ/ in *give* is more advanced in phrases with an abstract theme, such as *give a chance* than in its more concrete use of transferring an object, such as *give a pen*.

4.5.3. Schematic Slots as Exemplar Categories

Another role for exemplars in a construction's representation is seen in the way the schematic slots in constructions expand. A schematic slot in a construction might consist of a list of all the items that have occurred in that slot (as predicted by an exemplar model), or it might be considered a set of abstract semantic features that constrains the slot, as usually proposed. It could, of course, be both. However, the importance of the specific exemplars that have occurred in the construction can be seen when features that are redundant for particular items are referred to in an extension of the construction, as we will see below or in cases where a single abstract feature does not characterize a class or explain its extension.

For instance, it is often found that the set of lexical items that can occupy a slot in a construction may be constituted of two or more clusters of closely related items (Goldberg 1995; Israel 1996). An example is the well-known *way* construction which is used with verbs that indicate either manner of motion (4) or means (5) (see Jackendoff 1990).

(4) Romana Kryzanowska **weaves her way** through Drago's Gym in midtown Manhattan like a mother hen… (Time Magazine Corpus, 2004)
(5) His weakness, which Holmes unflinchingly describes, was an inability to resist the financial and sexual rewards that came along as he **clawed his way** to the top. (Time Magazine Corpus, 2000)

However, one also finds within these groups more specific clusters of verbs such as those indicating winding motion (*pick, thread, wind, wend, worm, weave*, etc.) or laborious motion (*plod, crawl, grind, slog, stumble,* etc.) and so on (Goldberg 1995; Israel 1996). Such clusters of highly related items are the results of item-based analogy (Skousen 1989; Eddington 2000; Bybee 2010: chapter 4). In other words, rather than making reference to a general semantic feature when using a construction, the speaker may very well reference a particular lexical item that has already been used in the construction and stored in memory. Exemplar-based representations of constructions will include a list of all the words experienced in a certain slot in a construction. This list of words (organized in clusters by similarity of meaning) is the basis for new extensions of the construction. Barðdal (2011*a*) shows that the majority of verbs borrowed into Icelandic take the case construction of synonymous verbs. Boas (2003) has argued for resultative constructions in English that many micro-constructions are available for reference when extending existing constructions to new lexical items.

Item-based extension leads to a family resemblance structure among the lexical items that constitute a schematic category in a construction. For instance, the set of adjectivals (including prepositional phrases) that occur in the Spanish 'become' construction using the verb *quedarse* can be analyzed into several clusters based on semantic similarity (Bybee and Eddington 2006). Some of these clusters are large enough that they include adjectivals that have no semantic relation to one another, but rather share properties with more central members in family resemblance fashion. Thus, our corpus study turned up *quedarse* with the adjective *quieto*, which means 'still.' It is, of course, related to *inmóvil* 'immobile' but also to an adjective with very different implications, *tranquilo* 'tranquil, peaceful.' This could in turn be related closely to adjectives expressing the more positive side of 'tranquility, satisfaction' with *conforme* 'satisfied' and this in turn to *a gusto* 'pleased' as shown in (6).

(6)

inmóvil	*quieto*	*tranquilo*	*conforme*	*a gusto*
'motionless'	'still'	'tranquil'	'satisfied'	'pleased'

As with other family resemblance chains, the ends of the chain may have very little to do with one another; *inmóvil* 'motionless' and *a gusto* 'pleased' share few features.

But they both share some features with *tranquilo* 'tranquil.' While *tranquilo* might have originally been used in this construction because of its similarity to *quieto*, it brought along with it the 'peacefulness' sense, which then could be referenced for further extensions. This type of structure, then, points to local extensions by item-based analogy, rather than global summations of abstract features. In order for item-based analogy or local similarity to apply, the cognitive representation of this construction must contain the specific adjectivals that have been experienced in it (see also Bybee 2010: chapter 5).

Local extensions may also be based on form, showing that lexical slots in constructions are not purely semantic, but contain information registered in memory for the form of items. The history of the Spanish 'become' construction with *quedarse,* as reported in Wilson (2009), shows that it first appeared with the adjective *solo* 'alone' in the thirteenth century, as well as with a few other adjectivals. The 'alone' sense gave rise in later centuries to uses with the adjectives *huérfano* 'orphaned' and *viudo* 'widowed,' as well as some phrases such as *sin heredero* 'without an heir' and *sin padre* 'without a father.' By the fifteenth century, there are many examples with *sin* 'without' but they extend well beyond the original semantics related to *solo* 'alone.' They include phrases such as *sin armas* 'without weapons,' *sin pluma* 'without a pen,' *sin deuda* 'without debt,' *sin pena* 'without grief,' and *sin quexa* 'without complaint.' The semantics of these phrases do not fit with the 'alone' semantics, nor do these phrases share abstract semantic features. All they have in common is the use of the preposition *sin* 'without.' This pattern suggests that features of form may also influence the choice of items occurring in a lexical slot in a construction. An exemplar representation would include these properties of the form, as all tokens and all their properties potentially have an impact on the representation. In contrast, if only semantic features are recorded in the construction's representation, there would be no way to account for this extension pattern.

4.6. Token and Type Frequency

4.6.1. Token Frequency

Another important argument for an exemplar model of linguistic representations is that exemplars, by their very nature, provide a record of the frequency of occurrence of tokens in linguistic experience. As mentioned above, each token of experience has an impact on the memory for linguistic items (Nader et al. 2000). In Bybee (1985), I proposed that representations are strengthened by repetition, making them easier to access. This proposal applies primarily to token frequency—the number of times a particular string occurs in a text or corpus. Thus, every time a construction is used, the constant parts of it are strengthened. But constructions also have schematic slots which will be represented by different exemplars

and which form categories as described in the preceding section. These categories can vary in their type frequency, that is, in the number of items that occur in the slot. Exemplar models also allow us to keep track of type frequency. Thus, the frequency profile of a construction can be quite complex. In this section we examine what is known about the token and type frequencies of constructions.

Let us take as our example the Resultative construction with *drive* consisting of an animate object with an adjective or prepositional phrase synonymous with *crazy*. This construction has recently begun to appear in hyperbolic uses in which the object is not literally insane, but only extremely irritated or distraught (see Bybee 2010: chapter 5). Here are some examples of different tokens of the construction:

(7) For the singers it was "up and down, up and down, from high C to low F," said Tenor Lawrence White. "It's enough to **drive you crazy**." (Time Magazine Corpus, 1965)
(8) "They won't let me out of their sight. It's **driving me crazy**." (Time Magazine Corpus, 1982)
(9) Though Lindsay finally lost him his good job and nearly **drove him wild** with anxiety about her morals, his fondness for her grew. (Time Magazine Corpus, 1933)

A construction is built up from such tokens with the sets of items that occur in each position forming categories. As with many constructions involving verbs and their complements, there is an open position that corresponds to subject position in other constructions. In most cases in this construction, it is inanimate, but it need not be. This would seem to be a truly open position and in that sense not defined by the construction. The object position (following *drive*) is occupied by animate objects, which can also be realized syntactically as pronouns.

The items that are specific to this construction are:

1. The verb *drive* occurs in different inflected and periphrastic forms, none of which are specific to this construction, yet they would all occur in exemplar representation of the construction, linked to the set of forms that *drive* has elsewhere. This construction occurs occasionally with *send* and *make* as the verb (Boas 2003).
2. The last position is occupied by a set of items synonymous with *crazy*. These include *mad, nuts, batty, wild, up the wall,* and *over the edge*.

A language user who has adultlike experience with English will have encountered a number of tokens of the construction, such as those in (7)–(9). All the tokens of this construction encountered in a corpus would constitute its token frequency. For instance, Boas (2003) finds 253 tokens of this construction with adjectives in the British National Corpus (BNC). One can also count the token frequency of a particular adjective in the construction, such as *wild*, which occurs twenty-two times in the BNC. Or we can be even more specific and count the number of times the exact form occurs in a corpus. For instance, *It drives me crazy* occurs thirty-three times in Corpus of Contemporary American English (COCA).[2]

What do these patterns of occurrence mean for cognitive representation? If each token of experience has an effect on representation, then each occurrence of the construction maps onto the exemplar cloud for the construction. The figure in (10) schematizes this exemplar cloud. Note that the particular items in the schema, such as *her* or *mad* are represented by clouds of exemplars themselves.

(10)

SUBJECT	[DRIVE]	*her*	wild
		him	mad
	[SEND]	*me*	*crazy*
		you	*up the wall*
	[MAKE]	*them*	*nuts*
		NP	*batty*
		. .	
		. .	

In (10) larger font is used to represent items that have a higher token frequency in the construction, that is, *DRIVE*, *me*, and *crazy*. We have shown in Bybee and Eddington (2006) that items with higher token frequency within the construction serve as the central members of the categories that form for schematic slots within the construction. That is, *drive* is the central verb for the construction and *crazy* (for American English) is the central adjective. Extensions of the construction will be based on these central members. In addition, these items of higher frequency will also be easier to access, which will increase their frequency even more.

As mentioned in section 4.4, constructions are formed by the domain-general process of chunking, by which repeated sequences of experience come to be remembered and processed as units. These processes operate gradually in language use, so that the creation of a construction as a unit does not mean that analyzability is lost immediately when constructions are formed. However, extensive repetition can lead to autonomy, as we will see in the next section. For the moment, note that a frequent instance of the construction, *it drives me crazy*, while it maps onto the schema in (10), also has its own sequential relations which may be stronger than other instances of the construction, such as *Lindsay nearly drove him wild with anxiety.*

4.6.2. Type Frequency

Type frequency is measured in the context of a construction and counts how many different items occur in the schematic slots of constructions. Thus, we might count the number of adjectives that occur in the *drive X crazy* construction in a corpus, as Boas did for the BNC, and find that there were eight (provided that *nuts* and *crackers* in this construction are considered adjectives). Counting the prepositional phrases would increase the type frequency of this slot. Constructional slots all have their own type frequency: the verb slot in (10) has a very low type frequency while

the subject slot is virtually unlimited within the domain of noun phrases and thus has a very high type frequency.

The importance of representing type frequency in the representation of a construction is that type frequency relates directly to productivity. In general, the higher the type frequency of a construction, the more likely it is to occur with a novel item (Bybee 1985; Baayen 1993; Barðdal 2008, 2011*a*). The effect of type frequency on productivity is constrained by two factors. First, items with very high token frequency may have formed a more autonomous chunk and do not activate the construction's exemplar cluster; thus such items will not contribute to productivity (Bybee 1985). Second, the semantic (and in some cases formal) properties of the schematic slot may restrict productivity. Thus, the complement to *drive* in (10) may include adjectives synonymous with *crazy* or phrases such as *over the edge, to distraction, to suicide,* but we do not find uses such as *drive X to pneumonia* or *drive X to fear.* In terms of formal properties, there is a construction *have the* NOUN *to,* where many of the nouns end in the suffix *-ity* or *-ability*, such as *have the capability to.* This perhaps accounts for the extension *have the availability to* (which seems only marginally acceptable to me, but was written on a sheet provided by an optometrist).

The semantic or formal properties of a slot determine its schematicity. A highly schematic slot, such as the object position in (10) has a broad set of properties, such as 'animate noun.' A less schematic slot might have a narrower definition, such as 'synonymous with *crazy*.' Schematicity and productivity are independent of one another: the '*crazy*' slot in (10) is fairly productive within this semantic domain. Thus, *drove me bananas, bonkers, berserk* are all possible extensions of this slot. Barðdal (2008) notes the inverse correlation between type frequency and semantic coherence, in that both properties encourage productivity, even though a very high type frequency often correlates with high schematicity and low schematicity (greater coherence) also gives productivity in the domain delimited. Despite the trade-off, the most productive categories have both high schematicity and high type frequency (e.g., the English regular verb formation with the suffix *-ed*). Both productivity and schematicity can be determined directly from an exemplar representation where all the types that have occurred in the construction are represented.

At least two factors contribute to the relation between high type frequency and productivity. First, the greater the number of types in a construction, the more bases there are for the item-based analogy that creates novel instances of the construction. Second, as Baayen (1993) points out, the processing of low frequency instances of constructions requires parsing or activation of the representation of the construction, which strengthens the construction. Processing high frequency instances of constructions, as we have already seen, can occur without activating the construction and therefore does not strengthen the constructions.

Another approach to frequency in constructions is Collostructional Analysis (Stefanowitsch and Gries 2003; Stefanowitsch, this volume). This approach rolls token and type frequency into one measure along with controls for overall frequency of the lexical item and a count of constructions in the language. In this

approach it is not possible to distinguish between the effects of token frequency and type frequency, which I find to be distinct effects and thus prefer to keep separate.

To conclude this section, both type and token frequency are important to our understanding of constructions as they affect category formation for slots in constructions, productivity of these slots, as well as the degrees of analyzability of constructions and particular exemplars of constructions. As frequency representations are an integral part of exemplar representation, these facts support exemplar models as a good choice for Construction Grammars.

4.7. How New Constructions Develop out of Old Constructions

A further important argument for the exemplar representation of constructions is the fact that new constructions develop out of existing constructions (Bybee 2003, 2006). In section 4.3 we discussed the famous *What's X doing Y?* construction and the fact that it has special pragmatic implications conventionally associated with it. The morphosyntax of this construction is completely unchanged from its source, which is a normal question with *what* and the progressive form of the verb *do,* yet it has entered the catalog of constructions of English, with a representation that is at least partially independent of the construction from which it arose. Exemplar representation is necessary to model the creation of new constructions: an instance of a construction with special implications is recorded in memory representation. If the same or similar instances occur with the same or similar implications, the exemplars will begin to form a cluster of representations such as that characteristic of a construction.

It is important to note that there is nothing in the meaning of a *what* question that can explain the implications of the new construction. Rather it is the specific contexts in which the construction is used that add the pragmatic implications. As I argued above, if these special implications found in the particular contexts were not made part of the representation for the instance of the construction, no new construction would arise. Thus, we see that speakers and hearers change language as they use it, and these changes are permanently registered in representation.

Once a new construction is formed, it can still be closely related to the construction from which it arose. However, over time and many usage events, a new construction can become more and more autonomous. This occurs via the chunking process discussed above. When a sequence is accessed directly without necessarily activating related lexical items and constructions, it strengthens the internal sequential relations, but does not strengthen any associations with other items (Hay 2001; Beckner and Bybee 2009; Bybee 2010). Over many repetitions of this process, the new construction can lose its analyzability. It is this process that occurs in grammaticalization. The *be going to* + VERB construction for future in English was

just one instance of a general construction for expressing movement in space for a purpose, but over the last few centuries it has lost its association with the purpose construction, as well as with the movement verb *go*. It has become an autonomous construction used for a distinct function, as in the example *A deal this big is going to carry a price tag* (Time Magazine Corpus, 2000), where the speaker is making a prediction and no movement in space is expressed. In addition, a grammaticalizing construction increases in type frequency as it comes to be used with more and more lexical items (Bybee 2003; Himmelmann 2004), as for instance, when the *be going to* construction comes to be used with inanimate subjects and stative verbs (see also Fried, this volume).

Thus, an exemplar model provides the representational basis upon which a particular instance of a construction can be established as a new construction and can become autonomous from its source.

4.8. Exemplar Semantics and Pragmatics for Constructions

The rich memory representation that is characteristic of exemplar models is compatible with the ideas expressed in cognitive linguistics that linguistic meaning is highly dependent upon and not separable from encyclopedic knowledge, or knowledge of one's world (Fillmore 1982, 1985*b*; Langacker 1987; Lakoff 1987; Croft and Cruse 2004). The representation of encyclopedic knowledge along with linguistic knowledge would emerge naturally from an exemplar model in which tokens of words or constructions are represented in memory along with the situations they have been associated with in experience.

Given the vast array of detail that memory can record, a pressing theoretical issue is how semantic categories are abstracted from the representation of experience and how abstract these categories really are. Note that this is not a question created in any way by exemplar theory. It is a question we must answer in any theory unless we want to postulate that linguistic categories are innately specified, for any theory must eventually address the nature of the abstraction process that takes place in language development in the child and the resulting relationship between exemplars derived from experience and abstract categories (see, for instance, Abbot-Smith and Tomasello 2006).

This question has been addressed for lexical meaning through research in prototype categorization, which has shown that the natural categories that humans construct are not based on highly abstract necessary and sufficient conditions, but rather consist of members that are more or less similar to a central member or prototype. The question concerning such categories is how the features on which similarity is based are chosen. Rosch and Mervis argue that prototype "categories form to maximize the information-rich clusters of attributes in the environment"

(Rosch and Mervis 1975*a*: 458). To this statement one might add, "as the environment is experienced by a human agent."

Grammatical meaning, as more abstract, is much less dependent upon encyclopedic knowledge, such as that represented in semantic frames (Fillmore 1982, 1985*b*). Grammatical meaning seems to make reference to the speech situation itself (as with deictic elements) or to internal relations within the utterance, by means of constructions. Yet, the meaning of constructions and grammatical morphemes must also be built up through specific tokens of use, so the question of the relation between the specific and the abstract still arises. Just as we must be able to identify attributes residing in different entities in the environment, and just as we are able to find similarities in the form of linguistic elements, so we are also able to find attributes that are shared by situations that we encounter in association with particular constructions. When shared attributes are identified, their representations are strengthened while the nonshared properties of situations are not reinforced. Thus, abstraction occurs because only certain features are associated with a construction.

The extent and nature of abstraction for grammatical meaning is still somewhat controversial, as it has been traditional to attempt to identify only one abstract feature for each grammatical morpheme (Jakobson 1971 and many others). Such analyses often run into difficulty, as we will see below. As exemplar models do not insist on the reduction of a category to only one feature, and as the research on natural categorization turns up prototype effects, it is important to examine the question of just how abstract grammatical meaning is. This question can be addressed through the natural experiment of language change, especially when viewed as ongoing. In the following I review what language change can tell us about grammatical meaning and argue that exemplar representation is necessary for our understanding of how these changes occur.

Just as recent research on idioms, prefabs, and other types of formulaic language (Nunberg et al. 1994; Wray 2002*a*; Corrigan et al. 2009) has emphasized the diversity and specificity of constructions and collocations in language use, so studies of grammaticalization have turned up many cases of uses of grammatical morphemes that are not as abstract as we were led to expect.

Under the heading of 'retention' (Bybee and Pagliuca 1987) or 'persistence' (Hopper 1991), we find examples in which the meaning or distribution of grammatical morphemes reflects their earlier patterns of use and meaning. For instance, while there is no question that English *will* expresses the abstract meaning of future, as in example (11), there are contexts in which a meaning of 'willingness' comes through as in (12) and (13), reflecting the earlier meaning of *will* as 'desire.' The second *will* in (13) can be interpreted as either prediction or willingness (negated to 'refusal').

11. Off of public lands in Nevada alone over the next five years, more than $10 billion worth of gold is going to be removed. And you and I *will* not get a penny for that. (COCA, 1990)

(12) This hacker is offering me $1,000 if I *will* give him a declined credit card that has a mother's maiden name (COCA Spoken, 2005)

(13) and now he got someone who *will* stand up and say, 'Well, this jury did a terrible job, because I know the case better, but gee, no one in law enforcement *will* listen to me.' (COCA, 1990)

Thus, despite the fact that grammaticalization yields an increasing abstractness or generalization of meaning leading from 'desire' to 'future,' in some contexts a more specific meaning is retained. These contexts are *if*-clauses as in (12), clauses with an indefinite subject, as in the first use in (13), and often in negative clauses, as in the second instance in (13).

In order for such a situation to be established and maintained over time, it cannot be the case that all tokens of *will* in its modal use are mapped onto one meaning. It must rather be the case when some uses of *will* were generalizing to express only future, exemplars of *will* in other contexts were unaffected, belonging to exemplar clusters partially independent from the more frequent, generalizing exemplars. This phenomenon is similar to the one we discussed earlier in connection with *what's X doing Y?*, whereby an exemplar of a construction takes on special implications and thus becomes a new, quasi-autonomous construction.

Another kind of example in which a very abstract grammatical morpheme is polysemous and thus would be represented by multiple exemplar clusters is cases where the meaning of the construction overrides the meaning of the individual elements in it. As the meaning of the construction is assigned to the whole construction, elements within it can lose their compositional meaning. For instance, *doing* in the *What's X doing Y?* construction is no longer interpreted compositionally. Similarly, with grammatical morphemes we sometimes find that their general meaning is obscured within certain constructions. For instance, when the English Past Tense occurs in an *if*-clause, it conveys the meaning 'hypothetical' rather than 'before the moment of speech.' Construction grammar matched with exemplar representation handles this fact nicely: the construction as a whole has a meaning and the occurrence of past tense used here is not mapped, except in form, onto the general semantic exemplar cluster for past tense.

These two examples show that despite the extreme abstractness of the meaning of future and past, there is no requirement that every token of these morphemes have the same abstract meaning. Rather autonomous exemplar clusters can develop for uses in particular constructions.

Another type of related case, also amenable to a construction cum exemplar account, occurs when it appears that the meaning of a grammatical morpheme or construction has been 'absorbed' from the contexts in which it is often used. The French negative construction of *ne* + VERB + *pas* is such a case. The noun *pas* 'step' has completely lost it earlier meaning in this context and in addition, it has absorbed the negative meaning from the construction so that it can be used in other constructions to indicate negative, as in *pas du tout* 'not at all' or *pas beaucoup* 'not much.'

The semantic prosodies found in corpus work (Sinclair 1996*b*) also demonstrate that exemplars of words or constructions are associated with contexts that can at once be very specific and cover long stretches of text. Smith and Nordquist (2012) argue that the association of negative affect with the *cause to* construction came about over time as the percentage of negative collocates increased. In Early Modern English, *cause* was as likely to occur with positive or neutral collocates as with negative ones, though the neutral collocates greatly outnumbered the positive ones, with the result that *cause* was associated with negative or neutral collocates. Smith and Nordquist (2012) find that in Present Day English (PDE), the negative collocates have come to dominate, representing 81% of the tokens they extracted from a PDE corpus. Note, however, that it is not the lexical item *cause* itself that has absorbed the negative affect, but rather the particular construction it occurs in. In the transitive construction, *cause* + NP (as in *cause an accident, cause damage*), a negative collocate is found in 90% of the cases, while in the infinitival construction (as in *cause the boat to sink*), the collocate is negative 60% of the time. In both cases, the other uses are neutral rather than positive. As might be expected, the general meaning of the construction is highly affected by the lexical types that occur in it, and it is the constructional unit that affects the meaning of its component parts.

These examples demonstrate that even though grammatical and constructional meaning can be very abstract in certain contexts, it is also the case that use in specific contexts can affect the meaning of grammatical and lexical morphemes, sometimes providing them with more specific meaning not found in all contexts. These specific meanings can be added in by inference, as we discussed in section 4.7, they can be retained from earlier uses in context, they can be absorbed from the overarching meaning of the construction they occur in, or they can be added by association with certain lexical items. All of these scenarios require cognitive representations in which a great deal of context is associated with individual tokens, which determines how they are mapped onto exemplars.

4.9. The Limits of Exemplar Models

An exemplar model is not, on its own, a theory of language; it is a model of memory representation. It addresses only the mechanism of memory storage and the resulting structures or categories and it says nothing about the content of linguistic categories, such as why languages tend to have past or future markers or categories such as 'subject.' Exemplar models speak to the degree of abstractness of memory representations, positing a rich, detailed memory for experience that spans many domains—linguistic form (phonetics), context, affect, and meaning. Much of the complexity needed for a complete model of language is not provided by the exemplar models that have developed for other domains. The levels of abstraction for constructions argued for in Croft (2001, 2003*a*) and Barðdal (2008, 2011*a*), for

instance, are not given by exemplar theory, but must be discovered empirically for linguistic constructions. Also, while exemplar models would allow exemplar clouds containing a word or other element that occurs in many different contexts, the factors that determine whether a word is analyzable in a certain context, such as *spite* in *in spite of*—that is, semantic transparency and frequency of use—are outside of exemplar theory.

At the same time one might argue that exemplar models are too unconstrained, as they allow any type of category to form and any property to provide the basis of a similarity matching. I do not consider this a problem, however, because it allows the substance of linguistic experience (as filtered through human cognition and attention) to be the prime determinant of the nature of the categories. Thus, it leaves many linguistic questions to be addressed empirically, through the comparison of languages and the study of language change.

4.10. Conclusion

Usage-based theory covers a wide range of research questions arising from the premise that use of language figures critically in determining the nature of cognitive representations of language, or put another way, usage events create linguistic structure. The premise of usage-based grammar leads directly to constructions with their direct association of meaning with form as the basic units of grammar. Exemplar representation is central to a usage-based approach to language, since it reflects the way that linguistic structure emerges when human cognition deals with tokens of experience with the world. I have argued here that exemplar representation helps us explain many of the dynamic properties of constructions, such as how they arise from other constructions and how they change over time.

A further important advantage of exemplar models is their ability to register token and type frequency for constructions. Frequency effects raise many interesting questions that are still under investigation, such as questions concerning the interaction of token frequency with type frequency, the gradual development of autonomy, the interaction of schematicity and type frequency in determining productivity, and the question of the effects of different levels of token frequency.

Two properties of the usage-based approach are worth emphasizing. First, representations are dynamic and change with usage events, not just across generations but within the individual as usage patterns change. The ability to represent change provides investigators with explanations for how and why constructions develop. Second, the processes we have examined here that create constructions and therefore grammar are domain-general. The processes that create constructions—cross-modal association (for linking sound and meaning or context), chunking, and categorization—are all processes that apply in other cognitive domains, such

as vision or motor production. These two properties together provide a linguistic theory with powerful explanatory possibilities.

NOTES

1. Time Magazine Corpus, http://corpus.byu.edu/time/.
2. Corpus of Contemporary American English, http://www.americancorpus.org.