19 Mechanisms of Change in Grammaticization: The Role of Frequency

JOAN BYBEE

One of the most notable characteristics of grammatical morphemes (hereafter "grams"; see Bybee and Dahl 1989) and the constructions in which they occur is their extremely high text frequency as compared to typical lexical morphemes. Since grams commonly develop from lexical morphemes during the process of grammaticization, one striking feature of this process is a dramatic frequency increase. This increase comes about as a result of an increase in the number and types of contexts in which the gram is appropriate. Frequency is not just a result of grammaticization, it is also a primary contributor to the process, an active force in instigating the changes that occur in grammaticization. This chapter treats two topics: (i) the manner in which the extreme frequency increase occurs, which will be examined via a case study of *can* in Old and Middle English; and (ii) those mechanisms of change associated with grammaticization that are attributable in some way to this dramatic frequency increase, including phonological, morphosyntactic, and semantic change. A third important theme of this chapter echoes that found in Traugott (this volume): none of these changes can be studied except in the context of the construction in which the grammaticizing element occurs.

1 The Grammaticization of Constructions

The recent literature on grammaticization seems to agree that it is not enough to define grammaticization as the process by which a lexical item becomes a grammatical morpheme, but rather it is important to say that this process occurs in the context of a particular construction (see Heine and Traugott, both this volume). In fact, it may be more accurate to say that a construction with particular lexical items in it becomes grammaticized, instead of saying that a lexical item becomes grammaticized. For instance, several movement verbs appropriately fit into the following constructional schema of English:

(1) [[movement verb + Progressive] + purpose clause (*to* + infinitive)]

E.g., *I am going to see the king*
I am traveling to see the king
I am riding to see the king

However, the only example of this construction that has grammaticized is the one with *go* in it. The particular example of this construction with *go* in it has undergone phonological, morphosyntactic, semantic, and pragmatic changes that have the effect of splitting the particular grammaticizing phrase off not only from other instances of *go* but also from other instances of this [movement verb + Progressive + purpose clause] construction.

2 The Role of Repetition

Also in the recent literature on grammaticization, we find extensive discussions of semantic change and its sources (see Heine et al. 1991; Traugott 1989; Bybee et al. 1994), but much less emphasis on the development of morphosyntactic and phonological properties of emerging grams. In an attempt to offer an integrated approach to the multiple changes that constitute grammaticization, I will focus in this chapter on the role that repetition plays in the various changes that a grammaticizing construction undergoes. The importance of repetition to grammaticization has been emphasized in Haiman's (1994) discussion of the parallels between the general cultural phenomenon of ritualization and the process of grammaticization in language, and in Boyland's (1996) examination of the effects of repetition on the cognitive representation of grammaticizing constructions. Building on these works, I will argue for a new definition of grammaticization, one which recognizes the crucial role of repetition in grammaticization and characterizes it as the process by which a frequently used sequence of words or morphemes becomes automated as a single processing unit.

Haiman (1994) makes a case for regarding the process of grammaticization as ritualization, citing the following aspects of ritualization, all of which are the result of repetition: *habituation* results from repetition and depletes a cultural object or practice of its force and often its original significance as well; repetition leads to the *automatization* of a sequence of units, and the reanalysis of the sequence as a single processing chunk, with formerly separate units losing their individual meaning; repetition also leads to the *reduction of form* through the weakening of the individual gestures comprising the act, and through the reorganization of a series of formerly separate gestures into one automated unit; and *emancipation* occurs as the original, more instrumental function of the practice gives way to a more symbolic function inferred from the context in which it occurs.

Applying these aspects of ritualization to the grammaticization process in particular, I will argue that frequent repetition plays an important role in the following changes that take place in grammaticization:

i Frequency of use leads to weakening of semantic force by habituation – the process by which an organism ceases to respond at the same level to a repeated stimulus (section 4).
ii Phonological changes of reduction and fusion of grammaticizing constructions are conditioned by their high frequency and their use in the portions of the utterance containing old or backgrounded information (section 5).
iii Increased frequency conditions a greater autonomy for a construction, which means that the individual components of the construction (such as *go, to* or *-ing* in the *be going to* example of (1)) weaken or lose their association with other instances of the same item (as the phrase reduces to *gonna*) (section 6).
iv The loss of semantic transparency accompanying the rift between the components of the grammaticizing construction and their lexical congeners allows the use of the phrase in new contexts with new pragmatic associations, leading to semantic change (section 7).
v Autonomy of a frequent phrase makes it more entrenched in the language and often conditions the preservation of otherwise obsolete morphosyntactic characteristics (section 8).

Before moving to an expanded discussion of each of these aspects of grammaticization, I will discuss the two ways of counting frequency in section 3, and demonstrate in section 4, with a case study of the development of *can* in English, how a grammaticizing construction increases its frequency.

3 Type and Token Frequency

Two methods of counting frequency are relevant for linguistic studies: one method yields token frequency and the other type frequency. Token or text frequency is the frequency of occurrence of a unit, usually a word or morpheme, in running text. For instance, *broke* (the past tense of *break*) occurs 66 times per million in Francis and Kucera (1982), while the past tense verb *damaged* occurs 5 times in the same corpus. The token frequency of *broke* is much higher than that of *damaged*. We can also count the token frequency of a grammaticizing construction, such as *be going to*, by counting just those occurrences of *be going to* that are used with a following verb (rather than a noun).

Type frequency refers to the dictionary frequency of a particular pattern, such as a stress pattern, an affix, etc. For instance, English past tense is expressed in several different ways, but the expression with the highest type frequency is the suffix *-ed*, as in *damaged*, which occurs on thousands of verbs. The pattern

found in *broke* has a much lower type frequency, occurring with only a handful of verbs (depending upon how you count them: *spoke, wrote, rode,* etc.).

The notion of type frequency can also be applied to grammaticizing constructions by counting the different lexical items with which a construction can be used: for instance, when in Shakespeare's English *be going to* had its literal meaning of a subject traveling to a location in order to do something, the subject position could only be occupied by a noun phrase denoting an animate, mobile entity, and the verb following the phrase would have to be a dynamic verb. As the phrase grammaticized and changed its meaning the number of different types appropriate for subject position expanded to include non-animate and non-mobile entities and the verb position expanded to include a broader range of predicates (e.g., current usage allows *The tree is going to lose its leaves; I'm going to be ready at nine*; etc.). A grammaticizing phrase is thus said to increase in generality (Bybee 1985) as the contexts in which it is appropriate move from very specific to more general.

A much-noted property of grammaticizing constructions is this increase in type frequency of co-occurring lexical items. As a consequence, the token frequency of units such as *going to* or *gonna* also increases dramatically. As important as the increase in type frequency or generality is, it is the high token frequency of grammaticizing phrases which provides the triggering device for many of the changes that occur in the form and function of the grammaticizing construction. High token frequency triggers many changes because it affects the nature of the cognitive representations in ways that will be explained as we proceed. First, however, we turn to the issue of the increase in token frequency of grammaticizing constructions, using the English modal auxiliary *can* as a case study.

4 How Does Frequency Increase? A Case Study of *can*

4.1 Generalization of meaning

One of the earliest-mentioned mechanisms of semantic change in grammaticization is bleaching or generalization, the process by which specific features of meaning are lost, with an associated increase in the contexts in which the gram may be appropriately used (Meillet 1912; Lehmann 1982).[1] In fact, generalization seems to characterize the entire grammaticization continuum – we note that as the process unfolds, grams always become more general and more abstract in their meaning, more widely applicable and more frequently used.[2] The mechanism behind bleaching is habituation: a stimulus loses its impact if it occurs very frequently.

Grammaticizing expressions have inherent meaning derivable from the meanings of their component parts. It is this inherent meaning that is said to be

Table 19.1 Stages of development for *can*

Stage	Meaning	
Mental ability	Mental	enabling conditions exist in the agent
Ability	_____	enabling conditions exist in the agent
Root possibility	_____	enabling conditions exist _____

Table 19.2 Contexts of use of *can*

Stage	Subject	Main verbs
Mental ability	Human agents	Intellectual states and activities Communicating Skills
Ability	Human agents	All of the above Overt actions and activities
Root possibility	Human agents Passive subjects Inanimate subjects	All of the above

bleached as grammaticization proceeds. In some cases (though certainly not all), a neat diagram may be constructed showing which parts of the original meaning are lost along the way. For instance, Modern English *can*, derived from an Old English main verb, *cunnan* 'to know,' can be charted as going through the stages in table 19.1[3] (cf. Bybee 1988 on *may*). At each stage, *can* is used in a wider range of contexts (table 19.2).

Ability and mental ability are self-explanatory; root possibility asserts that enabling conditions exist in general. They include the inherent abilities of the agent, but also factors in the external world that create enabling conditions. Examples follow:

(2) *Mental ability:*
Ful wys is he that kan hymselven knowe!
"Completely wise is one who knows himself!" (B. Mk. 3329)[4]

(3) *Skill:*
Ther seen men who kan juste and who kan ryde
"Men are seen there who can (i.e., know how to) joust and who can ride"
(A. Kn. 2604)

(4) *Ability:*
But I wol passe as lightly as I kan
"But I will pass by as lightly as I can" (B. NP. 4129)

(5) *Root possibility:*
Thou cannest not haue of Phocion a frende and a flaterer both to gether
"You cannot (It is not possible to) have of Phocion both a friend and a flatterer both" (UDALL Erasm. Apoph. 299a)

Tables 19.1 and 19.2 show what is meant by generalization or bleaching: specific features of meaning drop off, leaving a semantic core. The classes of main verbs with which the auxiliary *can* is used generalize, as does the range of possible subjects of *can*. However, this is not all there is to the story. It must be remembered that both specific and general meanings of a gram can coexist; old uses may be retained in certain contexts (Bybee and Pagliuca 1987; Hopper 1991). Furthermore, tables 19.1 and 19.2 are just schematic summaries; they do not actually inform us of how the changes took place. The result is generalization of meaning and contexts of use, but what exactly was the mechanism by which this generalization occurred?

4.2 *From noun phrase complement to verb phrase complement (Old English)*

The ancestor of the modern auxiliary *can* is the main verb *cunnan*, which expresses various types of knowing.[5] With a noun phrase complement denoting a person, a skill, or a language, the sense of knowing is acquaintance or acquired skill or knowledge (Goosens 1990). *Cunnan* is also used in the sense of understanding, as in "knowing the holy writings":

(6) Ge dweliað and ne cunnon halige gewritu
"You are led into error and do not know the holy writings" (Ags. Gospel of Matthew xxii)

In order for a main verb such as *cunnan* to begin its development into an auxiliary, it must expand its syntactic distribution to take verb phrase objects. *Cunnan* had very limited use with infinitive objects in the Old English period, so that studying the specific contexts in which it was used with an infinitive can give us some idea of how the development may have taken place. The infinitives used with *cunnan* in Old English mostly fit into three semantic classes of main verbs: verbs of mental state or activity, verbs of communication, and verbs describing skills. Table 19.3 shows the 13 examples listed in the OED of *cunnan* used with an infinitive before 1100, plus the additional items listed by Goosens from his sample.

Table 19.3 Verb classes used with *cunnan* in Old English

OED		Additional items listed by Goosens (1992)	
Mental states or activities:			
understandan		*geþencean*	'to comprehend'
ongietan (2)	'to understand'	*behabban*	'to comprehend'
tocnawan	'to distinguish, discern'	*wurðian*	'to esteem'
		gecnawan	'to perceive, know'
Communication:			
secgan (3)	'to say'	*sprecan*	'to speak'
geandettan	'to confess'	*tæcan*	'to teach'
		læran	'to teach'
Skills:			
gretan hearpan	'to touch a harp'		
huntian	'to hunt'		
wunda sniðan	'to cut a wound'		
Other:			
afandian	'to prove, try'		
bebeorgan	'to defend oneself'		

Goosens takes the mental state class as central and describes the other classes as related to this class more or less directly. There is no doubt that the mental state class is important, but when we consider how *cunnan* might have come to be used with infinitives, it seems likely that there were distinct motivations for the different semantic classes of verbs.

Since use of *cunnan* with mental state verbs is clearly important in Old English (and it continues to be important in Middle English), let us consider what would motivate the use of *cunnan* with verbs having a meaning that is so closely related to its own meaning. Indeed, meanings such as 'be able to know' and 'know how to understand' seem rather redundant. However, it is important to bear in mind that as a main verb, *cunnan* was already fairly frequent, and thus would have begun to lose some of its semantic force and specificity. I suggest that mental state infinitives then appear to be added in to bolster the meaning of *cunnan*, to flesh out the specific sort of knowledge intended in the context:[6]

(7) He ne con ongitan forhwy swylc God geþafað
he not *con* understand why such God allows
"He does not understand why God allows such as that" (950 Alfred's Boeth. xxxix)

The near synonymy of 'can't understand' and 'doesn't understand' supports the idea that *con ongitan* is a harmonic phrase which means about the same as either component alone would mean. Perhaps *cunnan* is beginning to bleach and grow too weak to stand alone in such contexts.

In other cases, it is not clear whether the form of *cunnan* means 'know how to' or is expressing a meaning similar to the main verb:

(8) Nu cunne ge tocnawan heofones hiw
Now *cunne* 2p distinguish-INF heaven-GEN hue
"Now you can distinguish/interpret heaven's hue" (Ags. Gospel of Matthew xvi.3)

In this passage the speaker is pointing out to the addressees that they know how to and do in fact interpret the color of the sky at sunset and dawn to predict the weather.

Thus it appears that one avenue by which *cunnan* begins to grammaticize as an auxiliary is determined by the fact that it was already frequent, and had already undergone some weakening of its semantic content. Of course, the use of *cunnan* with infinitives whose meaning is a more specific version than that covered by *cunnan* results in the further weakening of its semantic content.

With verbs of communication and instruction, *cunnan* is used in contexts in which it retains its 'knowledge' interpretation: it is not used in quotative contexts but rather where the content of what is said is asserted to be based on accurate knowledge of facts:

(9) þæt hi andsware secgan cunnen
that they answer say-INF *cunn*-PL
"That they can say the answer" (*c.*1000 Elena 374)

(10) Weras þa me soðlice secgan cunnon
man-PL then 1s-DAT truly say-INF *cunn*-PL
"Then men can truly say to me" (*c.*1000 Elena 317)

The third verb class used with *cunnan* in Old English consists of those denoting skills, particularly those with a strong intellectual component, such as reading, writing, or singing, but not excluding the more physical, such as hunting. This verb class corresponds to a set of nominal objects frequently used with *cunnan*. The most frequent nominal or pronominal objects of *cunnan* refer to people, but the second most frequent set comprises objects that refer to intellectual skills.[7] Thus we find 'know the holy writing,' 'know songs,' 'know book-learning,' 'know letters.' An infinitive construction could arise in these contexts by adding in the infinitive to an object that already consists of a noun phrase. Consider the following example, where the first instance of *cunnan* is followed by a noun phrase while the second has both a noun and an infinitive:

(11) ðy læs ðe him con leoða worn, oððe mid hondum con hearpan
unless 3s-DAT *con* song-PL many or with hand-PDat *con* harp
grētan, hafaþ him his gliwes giefe
touch-INF have-3s 3s-DAT 3s-GEN glee-GEN gift
"unless he knows many songs, or can (knows how to) touch the harp with his hands, has his gift of glee" (*c.*1000 Versus Gnom. 172)

In this case, it appears that the infinitive complement could develop directly out of the noun phrase complement as a vehicle for adding in more specific information about the skill being described.

The uses of these three verb classes in infinitival form with *cunnan*, then, appear to arise for different reasons, and perhaps are simultaneous developments. They are not necessarily totally independent, however. While each developing class of *cunnan* plus infinitive is a separate construction, it is plausible to assume that some more abstract generalization emerges from the similarities among these constructions.

Since the uses of *cunnan* are highly constrained lexically, they are appropriately described in a Cognitive Grammar (Langacker 1987) or associative Network framework (Bybee 1985, 1998) in which phrases or constructions are stored in the lexicon and generalizations are abstracted from these stored units. In this framework there is no strict separation of lexicon and grammar, but rather units of varying lengths and degrees of complexity may be stored lexically with the following properties: (i) the degree of strength or entrenchment of stored units is based on their text frequency; (ii) connections or associations of both a phonological and semantic nature are made among items, based on similarity or identity; and (iii) schemas of varying degrees of generality emerge from these representations.

A description of *cunnan* in Old English would require three quite specific schemas, one for each verb class, as shown in (12), and a more abstract schema, as in (13):

(12) a. *cunnan* + V-infinitive
'know' {mental state, activity}
b. *cunnan* + V-infinitive
'know to' {communicate, instruct}
c. *cunnan* + NP (V-infinitive)
'know how to' {skill} ({do a skill})

(13) *cunnan* + V-infinitive
'know (how) (to)' {activity involving mental capacity}

4.3 *Expansion to auxiliary status (Middle English)*

Since we are interested in how bleaching and generalization to new contexts take place, an appropriate time period to focus on is the Middle English period.

Once more our concern will be the semantic classes of verbs that appear in infinitival form with *can*. To determine the relative text frequency of the verb classes and individual members of these classes, it is necessary to examine all instances of *can* + infinitive in a stretch of text. For this purpose I have chosen the works of Geoffrey Chaucer and have examined the first 300 uses of *can* + infinitive listed in *A Concordance to the Complete Works of Geoffrey Chaucer* (Tatlock and Kennedy 1927), which includes all of the *Canterbury Tales*, most of *Troilus and Criseyde*, and several shorter poems.[8]

First we observe that the three verb classes that appeared with *cunnan* in Old English are still strongly represented in Middle English (in the following, the verbs are rendered in their Modern English spelling, unless that distorts the meaning or shape of the verb radically):

(14) Intellectual states or activities (52 tokens, 18 types):
deem, believe, see, know, guess, understand, *espy* (discover), judge, construe, imagine, comprehend, conclude, bethink, remember, find a difference, find a reason, shape a remedy, (wit) suffice

(15) Communication (102 tokens, 31 types):
clepen (name), *devyce* (describe), thank, say, tell (or count), express, expound, make mention, make a description, make by argument, answer, cry, bewail, speak, report, swear, lie, preach, *reherce* (describe), declare, reckon, amend, beguile, portray sorrow, assure, describe, write, complain, record, define, *distreyne* (urge), treat

(16) Skills ('know how to') (26 tokens, 18 types):
read, gloss, form, paint, *counterfete* (imitate), shape, do craft, do craftily, delve in herbs, work in philosophy, sing, dance, joust, play an instrument, play, entune, sound, make a beard

The frequency increase of *can* from Old to Middle English is manifested both in the use of *can* with a larger number of verbs of each class (i.e., type frequency) and in the development of a high token frequency for some combinations in the intellectual state and communication classes. Both kinds of frequency contribute to the bleaching of the meaning of an element.

Because of certain commonly used fixed phrases, the token-to-type ratio in the "intellectual states and activities" class and the "communication" class is approximately three to one. Here are the most commonly used main verbs:

(17) Communication class:

High token frequency:	tell	30
	say	29
	devyce (describe)	8
Type frequency:	31 distinct verbs	

(18) Intellectual states or activities:

High token frequency:	see	12
	deem	6
	understand	6
	espy (discover)	5
Type frequency:	18 distinct verbs	

In the associative Network or Cognitive framework described above, type frequency corresponds to the generality of the schema, which in turn corresponds to a higher degree of grammaticization. High token frequency corresponds to a local schema that is very strong or highly entrenched, such as *can say, can tell,* or *can see.* Increases in frequency of both types lead to the continued bleaching of the meaning of *can.*

Actually, the phrases listed above are abstractions from larger ritualized phrases found frequently in the Chaucer texts, phrases such as the following:

(19) I kan say yow no ferre (farther) (A. Kn. 2060)
I kan say you namoore (B. ML. 175; B. NP. 4159; G. CY. 651)

(20) more than I kan telle (B. ML. 1120)
mo than I kan make of mencioun (A. Kn. 1935)
more than I kan yow devyse (describe) (B. ML. 429)

(21) I kan nat seen (that) (B. Mel. 2735; TC II 794; TC IV 1365)

Each of these can be viewed as a construction with varying degrees of generality and varying degrees of entrenchment.

The Chaucer texts also reveal that the use of *can* with infinitives has expanded to other semantic classes of verbs, that is:

i verbs denoting states of mind that are not strictly intellectual, such as *love, suffer, have patience,* etc.;
ii verbs denoting states that are not mental or emotional: *be wrye* (twisted), *be rotten,* etc.;
iii verbs indicating a change of state in another person. These are probably related to verbs of instruction of Old English: *teach, heal, comfort, disturb,* etc.;
iv verbs indicating an overt action: *ride, go, send, climb, steal,* etc.

It is plausible to assume that these verb classes arose out of the earlier three classes gradually, since lines between semantic classes of verbs are not discrete (cf. the study of Kemmer 1995). I propose the following developments:

(22) 'know' > 'experience'
main predicates: Intellectual states > States of mind > States

(23) 'know to tell' > 'know how to' > 'be able to'
main predicates: Instruction > Change of state (transitive)

(24) 'know a skill' > 'be able to'
main predicates: Mental skills > Physical skills > Overt action

By the time these developments have occurred, there are very few predicates that cannot be used with *can*. Despite the generality with main predicates, *can* does not yet express root possibility with any regularity, since use with inanimate subjects is extremely rare: only 12 examples are found in the corpus of 300 and all but two of these are metonymic expressions for humans, that is, "inanimate" objects such as the eyes, the heart, wit, foolishness, and beauty. Two other inanimate objects that can tell or hide (the truth) are a book and the gossip or prattle (of women):

(25) As ferforth as my wit kan comprehend
"As far as my wit can comprehend" (TC IV 891)

(26) Swich vanyte ne kan don hire non ese
"Such foolishness cannot please her" (TC IV 703)

It appears that the most general schema for *can* in Chaucer's English is centered on human subjects and is only occasionally extended beyond humans to aspects of their behavior or metonymic uses of mind-body parts (such as eyes and wit). The most general schema, (27), does not have restrictions on the type of main predicate *can* occurs with. The tendency to use *can* commonly with certain semantic classes of verbs is captured in more specific schemas referring to the verb types listed in (22) through (24) or covering very specific constructions, such as those represented in (19) through (21):

(27) {human subject} *can* + infinitive
{know how to}
{experience}
{be able to}

At this period, *kan* has generalized to expressing human ability of all types, but it has not yet generalized beyond ability to include root possibility.[9]

4.4 *Further generalization: root possibility*

General ability is very closely related semantically and functionally to root possibility. While ability applies only to properties internal to an agent, root possibility includes both internal and external enabling conditions. It is paraphrasable as "it is possible for *x* to *y*." Thus in one of the few examples of

root possibility in the Chaucer texts, we can see how this paraphrase would apply:

(28) Til we be roten, kan we nat be rype
"Until we are rotten, it is not possible for us to be ripe" (A. Rv. 3875)

The close relation of ability to root possibility is due to practical considerations in the real world: the ability to do something often depends on both internal and external conditions. Thus in this example, again from Chaucer, it is difficult to tell if the conditions are internal, external, or both:

(29) Allas! kan they nat flee the fires heete? (G. CY. 1408)

Later in the Middle English period, examples interpretable as root possibility become much more common, and the syntactic conditions under which *can* is used continue to shift. In particular, the root possibility reading makes the use of *can* with stative predicates and in passive clauses possible, as the following two examples show:

(30) No worldely thyng can be wythout stryfe. (1509 Hawes Past. Pleas. xvi.xlix)

(31) Gij, But and thou array thy body sumptuously thou canst not be excused as chast in mind. (1540, Hyrde Vives' Instr. Chr. Wom. 1592)

Also examples of *can* expressing capacity, a use close to root possibility, appear in the sixteenth century:

(32) There is great number that fayne would aborde
our ship can holde no more. (Barclay Ship of Fooles 1570)

In this use, *can* begins to replace *may*, which, as we noted above (n. 9), was much used in the root sense in Chaucer's works. *May* is undergoing its own development, however, and beginning to be used more often in the epistemic sense ("it is possible that").

This brief survey of the development of *can* from Old to Middle English illustrates how the sharp frequency increase takes place: (i) the grammatical construction (*can* + infinitive) gradually extends to use with more and more types of verbs and then subjects; this extension is based on semantic similarity among the verbs in question, but its result is a generalization or bleaching of the meaning of *can*; (ii) certain phrases have a high token frequency, which also serves to bleach the meaning of their component parts. The result is a major change from the meaning of *cunnan*: in these root possibility readings of (30), (31), and even (32), no hint of the meaning of *cunnan* as 'know' remains.

5 Phonological Changes

A recognized concomitant of grammaticization is reduction in phonological form. In a large cross-linguistic sample, Bybee et al. (1991, 1994) demonstrate a significant association between degree of semantic grammaticization and phonological reduction, particularly in the length of the grams in question, but also in the degree of fusion of the gram with surrounding material.

The previous section illustrated in some detail the way increases in token and type frequency occur over time. In this section we will examine the link between frequency, phonological reduction, and fusion of grammaticizing phrases. The example of *can* is less useful here, since it is a single monosyllable, so other examples will be taken up. It should not be concluded, however, that *can* has undergone no phonological reduction just because its orthographic shape is fairly constant. Since the Old English period it has lost the final inflectional syllable that occurred in many forms (*cunnan, cann, canst, cunnon, cunne*) as have other verbs, and furthermore, in Modern English, *can* is phonetically reduced to [kŋ] or [ŋ] in high frequency contexts, such as after the pronoun *I*.

5.1 *Phonological reduction*

Recent studies of the lexical diffusion of regular sound changes have shown that in many cases, high frequency words undergo sound change at a faster rate than low frequency words. The effects of frequency have been shown for vowel reduction and deletion in English (Fidelholtz 1975; Hooper 1976a), and for the raising of /a/ to /o/ before nasals in Old English (Phillips 1980), for various changes in Ethiopian languages (Leslau 1969), for the weakening of stops in American English and vowel change in the Cologne dialect of German (Johnson 1983), for ongoing vowel changes in San Francisco English (Moonwomon 1992), for tensing of short *a* in Philadelphia (Labov 1994: 506–7), and for t/d-deletion in American English (Bybee 2001: 23ff, 112ff, 151ff).

Pagliuca and Mowrey (1987) argue that when one views articulation in terms of sets of overlapping gestures, all sound change can be classified as due to Substantive Reduction – the reduction in the magnitude of a gesture – or Temporal Reduction – the reduction in the duration of a constellation of gestures, resulting in the shortening of individual gestures or the increase in the overlap of gestures. This hypothesis is meant to explain the dominance of weakening and assimilation in attested sound changes. Browman and Goldstein (1990, 1992) make a very similar claim for casual speech processes (which I take to include the same range of phenomena as the category "sound change"). Browman and Goldstein hypothesize that all casual speech processes result from either the reduction in magnitude of a gesture, or the increase in the overlap of gestures.

These hypotheses await further investigation, but even if they turn out to have some counterexamples, the fact will remain that a large proportion of phonological changes are reductive in nature. Thus it is reasonable to ask why reductive changes would affect high frequency words or phrases earlier and at a faster rate than low frequency words and phrases. Several factors can be identified.

First, Fowler and Housum (1987) found that the second repetition of the same word in a single discourse was significantly shorter than the first token of the word. The speaker can be less explicit about the articulation of a word if it has already been used, because it will be easier for the listener to access if it has just been activated. Furthermore, Fowler and Housum point out that the reduction can actually be a signal to the listener that the word being used is just the same as one used earlier rather than a new and different word. It would follow then, that words or phrases that are often repeated in the same discourse (high frequency and grammaticizing phrases) would be in position to be shortened more often than words and phrases of low frequency.

Second, D'Introno and Sosa (1986) point out that frequency effects in the spread of a sound change are better viewed as familiarity effects: their position is that it is not so much the frequency of a word but rather its use in casual or familiar social situations that allows it to reduce or undergo change at a faster rate. Since the changes in question occur more often in casual speech, words that are used more often in casual speech will be more often subjected to the change.

Other factors might be involved as well, especially for grammaticizing constructions: as meaning generalizes, the informational contribution of the grammaticized elements decreases and along with that the intonational and rhythmic emphasis. Such prosodic reduction will have an effect on the segmental properties of the phrase as well.

For all of these reasons (and perhaps others), increasing frequency of use of grammaticizing constructions leads to phonological reduction. While the reduction is extreme in many cases, it usually follows patterns that are also seen in ongoing or future sound changes, suggesting that it is the frequency of use that hastens the changes. For instance, in Old Spanish, the second person plural suffix was *-des* (from Latin *-tis*), and was preceded by a stressed vowel: *-ádes*, *-édes*, or *-ídes*. In Old Spanish this medial *d* (pronounced [ð]) was gradually deleted, so that in Modern Spanish (in the dialects that use it) the forms are *-áis*, *-éis*, and *-ís*. Currently in most dialects of Spanish other instances of medial [ð] are gradually deleting. What is interesting is that this earlier morpheme-specific change was an instance of a more general change that would be current many centuries later.

Other instances of phonological reduction in grammaticization seem more extraordinary, but even most of these can be analyzed into steps that reflect the general reduction patterns of the language. For example, *going to* [goɪŋ tuw] reducing to *gonna* [gə̃nə̃] or even further, as in *I'm gonna* reducing to [aimə̃r̃ə̃], involves the following: (i) the reduction of full [o] to schwa; (ii) change of the velar to alveolar nasal; (iii) vowel nasalization; and (iv) flapping,

all of which occur in other words as well. On the other hand, certain aspects of this reduction are extraordinary: (i) reduction of the diphthong [oi]; (ii) flapping of [nt]; and (iii) deletion of [g] in [aimə̃ɾə̃].

5.2 *Phonological fusion*

Besides the reduction of the consonants and vowels within words, grammaticization often involves the phonological fusion of words or morphemes that formerly were separate. Here frequency is at work as well. Combinations of words and morphemes that occur together very frequently come to be stored and processed in one chunk. Boyland (1996) points out that as high frequency sequences of units come to be processed as single units, their gestural representation changes: what were previously multiple gestures come to be reorganized into single gestures and along with this reorganization comes reduction and increased overlap of gestures.

In Bybee and Scheibman (1999), we have shown that the reduction of the auxiliary *don't* in English is most extreme in precisely the phrases in which it most commonly occurs. Out of 138 occurrences of *don't* from spontaneous conversation, 87 occurred after the first singular pronoun *I*, making this the most common element to precede *don't*. There were 51 tokens in which the vowel was reduced to schwa and 50 of these occurred with *I*. (The other token was in the phrase *why don't you*, used to make a suggestion.) The reduction to schwa was also influenced by the following verb. The most common verb to follow *don't* was *know*, and 29 of the tokens with a schwa occurred with this verb. In fact, 29 out of 39 cases of *don't know* were reduced and all of these were in the phrase *I don't know*. The second most common verb to be used with *don't* was *think*, and 7 out of 19 of these cases, again all with *I*, were reduced to schwa. Other phrases in which *don't* was reduced were *I don't have (to)*, *I don't want*, *I don't like*, *I don't mean*, *I don't feel* and *I don't care*. The reduction did not occur with any other pronouns with the 20 other verb types found in the conversations.

We concluded that neither phonological nor syntactic conditioning is responsible for the reduction of *don't*, but rather that this reduction occurs inside of automated processing units, chunks that are automated primarily because they occur with high frequency. As *I don't know* comes to be produced as a single unit, the medial syllable loses its stress, allowing the vowel of *don't* to reduce.

6 Autonomy

Another consequence of a high frequency of use of a word or phrase consisting of multiple morphemes is a growing autonomy from other instances of these

same morphemes. Bybee (1985) argues that token frequency is an important determinant of semantic split among derivationally related words. That is, derived words that are of relatively high frequency (compared to their base form) are more likely to be semantically opaque and to have additional meanings or nuances not present in the base form. The reason for this is that high frequency words are present enough in the input to have strong representations of their own; they do not have to be understood in terms of other related words.

The same process applies to grammaticizing phrases – they gradually grow increasingly independent of their composite morphemes and other instances of the same construction. Thus the phrase *(be) going to* is becoming less and less associated with the individual morphemes, *go*, *ing*, and *to*, until a point may well come when speakers are surprised to find out what its etymological source is. Similarly, but on a different plane, *(be) going to* has disassociated itself from other instances of the construction, as given in (1). Such dissociations are phonological, semantic, and morphosyntactic.

Dissociations due to growing autonomy of grammaticizing phrases account for the splits that are often found between a morpheme in a grammaticizing phrase and its lexical source (Heine and Reh 1984; Hopper 1991). French *pas* in the negative phrase *ne . . . pas* is no longer associated with its etymological source, the noun *pas* meaning 'step.' The forms of *avoir* in French are still used for possession, but are also found in the construction of the Passé Composé (*j'ai chanté, tu as chanté, il a chanté*, etc.) and in the formation of the Future (*je chanterai, tu chanteras, il chantera*, etc.). In these three uses, despite similarities of phonology, these forms are best analyzed as autonomous from one another; they occur in different constructions and their meanings are in no way transparently related across these constructions.

7 New Pragmatic Associations

The autonomy of grammaticizing phrases and their growing opacity of internal structure makes it possible for new pragmatic functions to be assigned to them. Such new functions originate in the contexts in which the expressions are frequently used.

As an example, consider the phrase *I don't know* as used in colloquial American English. As mentioned above, this phrase can reduce to [aiɾ̃nə̃] or [aiə̃nə̃]. While it can be used with its literal meaning as an answer to a question, it can also be used in conversation to mitigate an assertion or to politely disagree or refuse something being offered (Scheibman n.d.). In these cases, *I don't know* is a single processing unit that is losing its association with the words from which it was derived. Due to its growing autonomy, it is capable of taking on new discourse functions that arise from the contexts in which it is commonly used.

8 Entrenchment: The Evolving Morphosyntactic Properties of English Auxiliaries

Another effect of high token frequency on complex forms is their maintenance of conservative structure despite the pressure of productive patterns (Bybee 1985). High token frequency explains why some English verbs (*ate, broke, wrote*) retain their irregular vowel changes despite the extreme productivity of the *-ed* affix for expressing past tense. High frequency constructions can also retain conservative morphosyntactic characteristics even in the face of new productive morphosyntactic patterns. Bybee and Thompson (1997) argue that even morphosyntactic constructions can exhibit this type of entrenchment due to the strength of the representation of the construction.

It is well known that English modal auxiliaries (*can, could, may, might, will, would, shall, should,* and *must*) have a set of syntactic properties that distinguish them from main verbs: the use of a bare infinitive, subject inversion in questions and other contexts, and the placement of the negative immediately following the auxiliary. How did these properties develop? Space is not available here for a detailed treatment of these properties, but the basic answer is that these properties were once variable properties of all verbs, but they have become conventionalized in these high frequency verbs, while all other verbs changed their properties in accordance with the changing syntax of the English language.

Consider first the use of the bare infinitive rather than the *to*-infinitive. In Old English, the infinitive was formed by adding *-an* to the verb stem. Thus verb + infinitive constructions in Old English had no intervening *to*. With general reduction of final syllables and the loss of inflections in verbs and nouns, the infinitive suffix gradually disappeared. Long before this suffix was lost, however, a new infinitive marker began to develop in the form of the preposition *to*. Haspelmath (1991) has shown that the primary source of infinitive markers cross-linguistically are allative or dative markers, which are first used in purpose clauses and subsequently generalized to other infinitival uses. This is exactly what happened in English: *to* with the infinitive (an erstwhile verbal noun) inflected in the dative was first used in purpose clauses and gradually extended to general use as an infinitive marker. During the Middle English period there was still some variation in the use of infinitives with and without *to*.[10]

Modals such as *can* have very consistently occurred throughout their history in constructions without *to*. The reason for this is that these constructions were first created and apparently entrenched before *to* developed as the infinitive marker. Since constructions with the modal auxiliaries were of high frequency and thus highly entrenched, they were not reformulated after the *to*-infinitive generalized in the language. The same is true of other verb + infinitive constructions that have survived from the Old English period. For instance, *go* +

infinitive and *see* + infinitive constructions use bare infinitives even today: *Let's go see; I saw him do it*. More recent formations with functions similar to those of the modal auxiliaries, such as *want to, be going to, have to,* use the newer infinitive construction that was established before these constructions became entrenched. Thus it is the fact that the constructions with *can* + infinitive arose before the *to*-infinitive and the fact that they were of high frequency that together explain why *can* uses a bare infinitive.

Another striking characteristic of the class of auxiliaries to which *can* belongs today is that they invert with the subject in certain constructions, primarily questions, but also (perhaps archaically) in conditional protases lacking *if*, and in clauses with fronted negative elements. In the Middle English period this verb–subject order in these contexts was a variable property of all verbs; it was not restricted to auxiliaries (Mossé 1952: 126–8). Consider these examples:

(33) Gaf ye the chyld any thyng?
"Did you give the child anything?"

(34) Ne sunge ich hom never so longe,
Mi songe were i-spild ech del
"Even if I sang to them ever so long, My song would be entirely lost (on them)"

Since the modal auxiliaries and *be* and *have* as auxiliaries were becoming increasingly frequent in this period, they would commonly occur before the subject in these contexts. While other verbs eventually ceased to appear in this position, taking instead the position after the subject which eventually became obligatory, the auxiliaries, including the newly developed pro-verb, *do*, remained in inverted positions in these special constructions. Again it is their high frequency that accounts for their conservative behavior. The constructions with inverted auxiliaries were highly entrenched and thus not prone to revision despite the other syntactic changes occurring in English.

The position of the negative *not* after *can* and other auxiliaries has a similar diachronic explanation. The sentence negation particle in Old English, *ne*, occurred before the verb, but in Middle English, it was reinforced by another negative *nought, not*, which derived from *ne* + *wiht* (literally: 'not a creature'). *Not* occurred after the verb in Middle English and became the normal negative marker as the preposed *ne* was lost. It occurred after simple finite main verbs as well as after the auxiliary (Mossé 1952: 112):

(35) My wyfe rose nott

(36) cry not so

The position of the negative after *can* and other auxiliaries is the preservation for this high frequency group of the order that once applied to all verbs. While

other verbs require the use of *do*-support, the auxiliaries have simply continued to participate in the highly entrenched construction that was established in the fourteenth century.

Thus it can be said that the special properties of the auxiliaries in English are the retention of older morphosyntactic properties that were once general to English verbs. These modal auxiliaries and the other auxiliaries, *be, have,* and *do,* have retained these properties because of their high frequency: due to repetition their participation in certain constructions is highly entrenched and not likely to change. By the same token, modal constructions developing more recently will reflect the morphosyntax of the period in which they develop and are highly unlikely to fall in with the older modals and take on their characteristics, such as using the bare infinitive and occurring before the subject in questions.

This preservation of older morphosyntactic characteristics in high frequency constructions can be attributed to the same mechanism as the preservation of irregularities (older morphological properties) in inflected forms. While analogical change generally operates to level or regularize morphophonemic alternations (e.g., as *wept* becomes *weeped*), forms with high token frequency tend to resist such change (e.g., *kept* is not becoming *keeped*); see Anttila, Dressler, and Hock, all this volume).

9 The Effects of Repetition

This survey of the changes that occur in grammaticization has revealed that repetition affects semantics and phonology by promoting change, in particular, reductive change, and that repetition affects morphosyntax by ensuring the retention of older characteristics. It might seem contradictory that repetition could both encourage innovation in one domain and enhance conservatism in another. This paradox is also found in the lexical diffusion of phonetic versus morphophonemic change. In Hooper (1976a), I pointed out that sound change affects high frequency items first, while analogical leveling affects low frequency items first. The substantive properties of words or phrases, their meaning and phonetic shape, are modified, usually reduced, with use. The ritualization or automatization process has an on-line effect of compressing and reducing; this is a processing effect. In contrast, the structural properties of words and phrases – that is, the morphological structure of words and the syntactic properties of constructions – are preserved by repetition; this is a storage effect. Frequently used words and phrases are highly entrenched and more likely to be accessed as whole units and less likely to be reformed on-line. Thus their general structure – the morphological irregularity of high frequency nouns and verbs, or the structure of high frequency constructions – will tend to be preserved. We can say, then, that repetition has a reductive effect on-line, but a conserving effect in storage.

Repetition is universal to the grammaticization process. Repetition and its consequences for cognitive representation are major factors in the creation of grammar. The conventionalized aspects of language provide the framework for manipulation of our thoughts into objects of communication. Repetition alone, however, cannot account for the universals of grammaticization. The fact that the same paths of change are followed in unrelated languages has multiple causes. It is not just the fact of repetition that is important, but in addition *what* is repeated that determines the universal paths. The explanation for the content of what is repeated requires reference to the kinds of things human beings talk about and the way they choose to structure their communications.

ACKNOWLEDGMENT

The author wishes to thank John Haiman, Barbara Need, Sandra Thompson, and Elizabeth Traugott for helpful comments on an earlier draft of this chapter.

NOTES

1 Care must be taken here to distinguish between meaning and use: as a gram loses specific features of meaning, it appears to take on more uses. Being used in a wide range of contexts does not mean that the gram has *more* inherent meaning.

2 At the end of the grammaticization process, an old gram may be restricted in use by newer grams that replace some of the uses of the older one. The consequent addition of contextual meanings to the old gram may appear to make meanings more restrictive.

3 The permission use of *can* is not treated here. In Bybee and Pagliuca (1985) and in Bybee et al. (1994), we argue that the permission use of grams originally expressing ability develops out of the root possibility sense. Root possibility expresses a highly generalized set of enabling conditions, which include the social conditions that govern permission.

4 Abbreviations for examples from Chaucer are: B. Mk. (Monk's Tale); A. Kn. (Knight's Tale); B. NP. (Nun's Priest's Tale); B. ML. (Man of Law's Tale); G. CY. (Canon's Yeoman's Tale); B. Mel. (Tale of Melibeus); TC II (Troilus and Criseyde, book 2); TC IV (Troilus and Criseyde, book 4); A. Rv. (Reeve's Tale). All other abbreviations are from the OED and follow the OED's format for dates and details for locating the example in the text: date of publication, author/title of work, chapter, page number, etc.

5 The Past Tense of *cunnan*, OE *cuþ*, which gives Modern English *could*, will not be treated here. See Bybee (1995) for the development of Past Tense modals in English.

6 Lyons (1977) refers to cases in which two modals of similar meaning co-occur in a clause

without increasing or decreasing the degree of modality as modal harmony. *Cunnan* 'to know' plus a verb of knowledge could be regarded as an harmonic expression.

7 This is true of Goosen's sample and the small OED sample consulted.

8 There are several varieties subsumed under Old English and even more under the designation Middle English, so it cannot be assumed that there is necessarily a direct developmental relation between the languages represented in the texts used here. Still it is clear that in some general sense a type of diachronic relation exists.

9 In Chaucer's English, root possibility is expressed by *may*, which derived from a verb expressing physical power or ability. *May* is more grammaticized semantically than *can*: in the Middle English period it is used frequently with inanimate and generic subjects to express root possibility. It is also commonly used in subordinate clauses and is even beginning to express epistemic possibility in some contexts (see Bybee 1988).

10 Indeed in Modern English there is still variation between the bare infinitive and the infinitive with *to*, as in *help someone (to) do something*.