
Pragmatic factors

4.1 Introduction

Although it is possible to describe change in terms of the operation of successive strategies of reanalysis (rule change) and analogy (rule generalization), the important question remains why these strategies come about – in other words, what enables the mechanisms we have outlined, most especially those involved in grammaticalization.¹ It is tempting to think in terms of “causes” and even of “explanations” in the sense of “predictions.” However, the phenomena that give rise to language change are so complex that they will perhaps never be understood in enough detail for us to state precisely why a specific change occurred in the past or to predict when one will occur and if it does what it will be (Lass 1980). Rather than referring to “causes” or “explanations,” we speak more cautiously of motivations or enabling factors, understanding always that we are referring to potential and statistically preferred, not absolute, factors (see, among many others, Greenberg 1978b; Romaine 1982; Croft 2000; Maslova 2000).

As mentioned previously, among motivations for change three have been widely discussed in recent years. Of greatest interest within generative linguistics has been the role of language acquisition, especially child language acquisition. Sociolinguists, by contrast, have tended to focus attention on the role of communities and different types of contact within them. Of special interest to those working on grammaticalization has been the role of speakers and hearers negotiating meaning in communicative situations.

Here we put forward arguments for the view that there are a number of competing motivations which can all in some sense be said to be examples of maximization of economy or “simplicity”: basically they can be summarized as maximization of efficiency via minimal differentiation on the one hand, and maximization of informativeness on the other. On this view, hearers play a major role in change because they process input in ways that may not match the speaker’s intentions. But speakers also play a major role in enabling change, because in producing speech they have communication as their goal, and therefore are always in search of ways to guide the hearer in interpretation. In an ideal communicative situation,

speakers take responsibility for success in communication and seek to meet hearers' attempts "to integrate new information with information that is already accessible" (Blakemore 1990). However, differences in what is actually accessible in the communicative situation based on differences in age, social background, culture, attention, or other factors may over time cumulatively lead to change. Furthermore, the motivations of simplicity and informativeness are inevitably in competition in the individual language user (see, e.g., Du Bois 1985; Bates and MacWhinney 1989), and therefore the development of language involves conflict and problem solving (Dressler 1985).

There are great difficulties in defining the notion of economy in anything like rigorous terms. We know very little about what does and does not take "effort" in producing or interpreting utterances, and still less about what would constitute economy of mental effort on either speaker's or hearer's behalf, although we probably know more about simplicity of perception than of production. Nonetheless, there seem to be useful, if sometimes intuitive, notions involved.

In considering the hearer's role, it is usually assumed that the hearer will seek the most unambiguous interpretation, and furthermore that the hearer is actively engaged in interpreting (usually abducing) input, whether as a child or an adult in the process of language acquisition. It is therefore not surprising that work on hearer motivations focuses on perception, and on meanings interpreted from the linguistic data which is the input to the acquisition process. However, it is also possible to think of hearers as the targets of speakers' output. From this perspective, hearers motivate speakers' intent to be informative and clear.

In considering the speaker's role, it has been customary to think of the tendency to reduce the speech signal, e.g., via rapid speech, a process resulting in "signal simplicity." The reduction of *be going to* to *be gonna* is an example. Signal simplification typically results from the routinization (idiomatization) of expressions. Rather than find different ways of saying approximately the same thing, speakers will repeat expressions they have heard before, even if they are in competition with other expressions, perhaps in the interest of sounding "with it." Well-known recent examples include the use of *you know*, *be all*, etc. Such routinized, or idiomatized, expressions can be stored and used as simple units. Naturally they are more frequent in discourse than expressions created and used "on the fly," which may indeed be novel, once.

Idiomatization of expressions tends to lead to reduction and simplification of the signal. With this process in mind, Langacker has said: "It would not be entirely inappropriate to regard languages in their diachronic aspect as gigantic expression-compacting machines" (1977: 106). However, compacting, obliteration of boundaries, and reduction of redundancy is balanced in normal language situations by the introduction of new and innovative ways of saying approximately the same thing.

These new and innovative ways of saying things are brought about by speakers seeking to enhance expressivity. This is typically done through the “deroutinizing” of constructions, in other words, through finding new ways to say old things. Expressivity serves the dual function of improving informativeness for the hearer and at the same time allowing the speaker to convey attitudes toward the situation, including the speech situation. This very process of innovation is itself typically based on a principle of economy, specifically the economy of reusing extant forms for new purposes (Werner and Kaplan 1963; Slobin 1977). To return to our example of the extension of *be going to* to the intentional future: the directional phrase has been reused; it is more substantive (phonologically longer) and therefore more accessible to hearers than, e.g., *'ll* or even *will*. As a future it is also more based in the speaker’s subjective attitude and perspective on what is being talked about than is its locational counterpart.

While the competition between motivations has been a major topic of research, individual researchers have tended to be interested in different subspects of the complex balance between creativity on the one hand and routinization on the other. Therefore studies of grammaticalization have taken different paths. A focus on hearers’ tendency to reanalyze abductively may correlate with work on morphosyntactic changes in grammaticalization (e.g., Langacker 1977). A focus on speakers’ expressive use of language to get a point across may correlate with work on lexical origins of grammaticalization, especially on pragmatic enrichment of lexical items in the early stages of grammaticalization (e.g., Heine, Claudi, and Hünemeyer 1991a,b; Traugott and König 1991). A focus on speakers’ tendency to economize may correlate with work on routinization, and frequency (e.g., Givón 1979, 1991b; Hopper 1987; Krug 1998, 2001; Bybee and Scheibman 1999; Bybee 2002). These approaches are, of course, not necessarily contradictory, although they do sometimes lead to different views of what should or should not be considered a case of grammaticalization. Any comprehensive study of a particular set of grammaticalization changes that lasts over long periods of time and involves continued grammaticalization must ideally take cognizance of all three kinds of approach.

Before concluding this section, we should note that the claim that grammaticalization (and indeed language change in general) is motivated by speaker–hearer interactions and communicative strategies is a claim that change is goal-directed. For example, Heine, Claudi, and Hünemeyer claim that “grammaticalization can be interpreted as the result of a process that has problem solving as its main goal” (1991a: 29). There has been a tendency in American linguistics to distrust arguments about goal-directedness, partly on the grounds that they cannot be empirically proven, partly because the important research question has until recently been considered to be how language is understood out of context. Arguments

based on mechanisms for change, and on (passive) capacities for language, have been privileged over arguments based on purposes to which language is put. For example, Bybee has argued that the development of morphology is spurred on by spontaneous processes whereby semantic functions that are similar are subsumed under closely related grammatical functions (Bybee 1985: 204; see also Bybee, Perkins, and Pagliuca 1994: 297–300). These processes are in Bybee’s view cognitive rather than communicative.

Part of the problem with the concept of goal-orientedness is that it is often discussed in terms of “need” or set goals, in other words, in “teleological” terms. Clearly “communicative need” is not a plausible motivation in most cases of grammaticalization, since not all languages express the same grammatical functions, and even less do they express the same grammatical functions in the same way. Furthermore, to assume that as an older system becomes eroded it may cease to function at an adequate level of communicative coherence and therefore must necessarily be revitalized (as suggested by Givón 1982: 117, cited in Bybee 1985: 202) is to posit a stage of language such as is unknown, in other words it violates the uniformitarian principle and is not empirically supported by the data. Most emphatically, languages are not goal-oriented. We agree that “the view which ascribes language a will of its own, a sort of conscious control over its own future, seems to us gratuitous and untenable. It remains true, however, that language is a communicative tool at the disposal of its speakers, to whom the attribution of an independent will and volition is considerably less controversial” (Vincent 1978: 414). The position we take is that users may be consciously or unconsciously goal-oriented (see Keller 1994 for a detailed account of goal-orientedness in language change that is not teleological in the sense mentioned above). In speaking of communicative strategies and problem solving in the course of speaker–hearer interaction, we refer not to filling gaps, but rather to strategies used by speakers and hearers in producing and understanding the flow of speech as it is created. These speakers have intentions, and their intentions may lead to change over time. However, barring cases of “language engineering,” e.g., attempts to legislate against use of the derivative *-man*, or exclusive use of *he* as a generic pronoun, speakers for the most part do not intend to change the language. On the contrary, many would like to prevent change if possible.

4.2 Inferencing and meaning change

The discussion in the preceding section has made reference to various motivations such as economy, efficiency, clarity, expressivity, and routinization. Such motivations are issues of usage and speaker–hearer purposes, and can be

called “pragmatic,” that is, they have to do with the relationship between language and the contexts in which it is used, including other instances of language, most especially the meanings that arise from this relationship (Levinson 1983, 2000).

Since pragmatics by definition deals with meanings beyond structure, many linguists working in the tradition of formal grammars for a long time excluded pragmatics from consideration in accounting for motivations for change. Many also excluded the meanings usually treated by lexical semantics, that is, components of sense and their relation to the objects to which they refer. For example, in discussing the syntactic changes involved in the development of the modals, Lightfoot (1979) argued that syntactic change was autonomous, i.e., independent of semantic or pragmatic motivations. He attempted to show that the modals “underwent very many changes in their syntax and in their meaning but . . . these changes seem to have proceeded quite independently of each other” (1979: 100). Indeed, he went on to say: “it does not seem possible to define a class of modals (and therefore of preterite-presents) on semantic grounds” (1979: 103). Such an approach is a natural consequence of conceptualizing significant change as resulting from children’s attempts to seek cues to grammar as a computational system (see, e.g., Lightfoot 1999), rather than as a strategic one.

The approach from “autonomous syntax” has been called into question by many linguists working on grammaticalization. This follows inevitably from the interest in changes whereby a lexical item becomes a grammatical one, because a lexical item by definition has semantic as well as syntactic, morphological, and phonological properties. But it also follows from the concern that a theory which regards semantic change as independent of morphosyntactic change provides no reasoned account for the extensive evidence that grammaticalization affects similar classes of lexical items in similar ways across a wide number of languages. This point was made especially cogently by Bybee and her colleagues in various studies of verbal morphology (e.g., Bybee 1985; Bybee and Dahl 1989) and by Heine and his colleagues in various studies of nominal morphology (e.g., Heine, Claudi, and Hünemeyer 1991a); valuable resources for putative examples (not all with empirical historical support) of cross-linguistic and cross-categorial grammaticalization include Heine *et al.* (1993) and Heine and Kuteva (2002).

Indeed, much work on grammaticalization since the early 1980s has focused extensively on the kinds of meaning changes involved in grammaticalization and the cognitive motivations behind them. For some, the meaning changes are regarded as semantic, e.g., “semantic change leads to the development of grammatical meaning” (Bybee and Pagliuca 1985: 59). For others it is mainly pragmatic (Traugott and König 1991). For some it is thought to be motivated primarily by metaphorical processes (Claudi and Heine 1986; Sweetser 1990), for others it is thought to be motivated by associative or “metonymic” as well as metaphorical processes

(e.g., Traugott and König 1991; Heine, Claudi, and Hünemeyer 1991a). The position we take here and will elaborate on in subsequent chapters is that meaning changes and the cognitive strategies that motivate them are central in the early stages of grammaticalization and are crucially linked to expressivity. Furthermore, the meaning changes are initially pragmatic and associative, arising in the context of the flow of speech. At later stages, as grammaticalization continues and forms become routinized, meaning loss or “bleaching” typically occurs, but even so, older meanings may still continue to constrain newer, “emptier” ones.

Before we proceed, a few words about some of the widely accepted ideas concerning semantics and pragmatics may be helpful.

4.2.1 *Semantics and pragmatics*

There is a vast literature on semantics and pragmatics but as yet little consensus on exactly where the boundaries between the two areas lie, or even whether there are indeed boundaries. Nevertheless, there is a pre-theoretical sense in which it is clear that a distinction needs to be made between the sentence (semantic) meaning of *Can you pass the salt?* (= ‘Are you able/willing to pass the salt?’), the expected response to which would be *Yes* or *No*, and the utterance (pragmatic) meaning (= ‘Please pass the salt’), the expected response to which is the non-linguistic action of passing the salt. For purposes of this chapter, it must suffice to note that we believe that linguistic theory should eventually provide an integrated account of semantics and pragmatics.

It is useful to distinguish between that part of semantics that concerns lexical, phrase, and sentence meaning, and that part of pragmatics that concerns inferences about linguistic meaning based on contextual assumptions such as the cooperativeness of participants in a conversation. The view of the relationship between semantics and pragmatics we adopt here is as follows (Levinson 1983, 2000; Green 1996 [1989]). Semantics is primarily concerned with meanings that are relatively stable out of context, typically arbitrary, and analyzable in terms of the logical conditions under which they would be true. Pragmatics, by contrast, is primarily concerned with the beliefs and inferences about the nature of the assumptions made by participants and the purposes for which utterances are used in the context of communicative language use. It concerns both speakers’ indirect meaning, beyond what is said, and also hearers’ interpretations, which tend to enrich what is said in order to interpret it as relevant to the context of discourse (Sperber and Wilson 1995[1986]). Many of the beliefs and inferences which are the subject of pragmatics are thought to be cognitively universal, in other words, not learned or arbitrary. Furthermore, they are rarely if ever subject to analysis in terms of truth.

4.2.2 Relationships between senses of a form: homonymy and polysemy

An issue on which there is little agreement across various subfields of linguistics is exactly how to characterize the relationship between the various senses of a form. One approach is to maximize difference and homonymy (same form, unconnected different meanings). Rejecting the traditional literary practice of grouping together all related meanings that can be associated with a single phonological form, McCawley (1968: 126) suggested that there is no a priori reason for grouping items together in a dictionary: one could take the notion “lexical item” to mean the combination of a single semantic reading with a single underlying phonological shape, a single syntactic category, and a single set of specifications of exceptional behavior with respect to rules. He argued that *sad* in the two meanings displayed in (1) should be analyzed as two separate items:

- (1) Sad₁: ‘experiencing sadness, said of a living being’
Sad₂: ‘evoking sadness, said of an esthetic object’

Likewise, from this perspective the *can* of ability and possibility and permission are unconnected semantically. Another approach is to maximize similarity and monosemy (same form, same meaning). For example, Groefsema (1995) argues that the various meanings of the modals are so indeterminate that they should be analyzed as sharing a unitary underspecified meaning; from this perspective, there is one *can*. A third approach is to argue that certain forms share conceptually related meanings (polysemy) (e.g., G. Lakoff and Johnson 1980; Sweetser 1990). Bybee, for example, suggests that there is one form *can* in PDE with related meanings:

- (2) a. Mental ability: e.g., I can read German.
b. Physical ability: e.g., I can swim a mile.
c. Root possibility: e.g., This word can be used in many contexts.
d. Permission: e.g., I can take books out for two weeks. (Bybee 1988: 256)

It is often argued that the fine, sometimes minimally discrete, meaning distinctions between various stages of grammaticalization or between focal clusters on a cline call for a theory in which different meanings may be closer or more distant (see G. Lakoff 1987 on the concept of “networks” of polysemies; and, from a different perspective, Kemmer 1993b on semantic maps of related terms within the domain of voice). For example, mental and physical ability are more closely related to each other than to permission among the polysemies of *can*.

In general, from the perspective of grammaticalization it is methodologically essential to assume polysemy if there is a plausible semantic relationship, whether or not the forms belong to the same syntactic category, because otherwise relationships between more and less grammaticalized variants of the same form cannot be

established, either diachronically or synchronically. What constitutes a “plausible semantic relationship” has until fairly recently been a matter of considerable debate. Drawing on Haiman, especially (1985a), Croft (1990) discusses a useful heuristic for the distinction between the homonymy of *two*, *to*, *too* (all [tu]), and the polysemy of the directional and recipient meanings of *to* in *I drove to Chicago*, *I told the story to my brother*. He suggests that a major criterion is evidence from cross-typological comparison: “if many diverse languages independently have the same pattern of ‘homonymy,’ then the meanings are closely related” (p. 166). *Two*, *to*, *too* do not tend to be expressed by the same form cross-linguistically; this, in addition to their lack of historical relatedness, provides evidence that they are homonymous in English. However, directional and recipient *to* are frequently, indeed typically, expressed by the same form cross-linguistically (cf. also *-lle* in Finnish). This, together with their historical relatedness, provides evidence that they are polysemous.

There has been a long history in linguistics of concern about the notion of polysemy. For example, Bolinger has said: “the natural condition of language is to preserve one form for one meaning, and one meaning for one form” (Bolinger 1977: x, cited in Haiman 1985a: 21). This concern presumably stems from the “scientific” approach to language which is the foundation of linguistics. To oversimplify, from this point of view the optimal language would be one in which every meaning was distinct, just as every numeral is distinct (the “idealized language” we referred to in Section 3.3). However, such “optimality” would clearly in actual fact be dysfunctional since there are far too many meanings for the brain to remember individual expressions for them. “One form – one meaning” is an ideal on the dimensions of choice of form and the motivation to maximize information. It is balanced and offset by another optimality, that of associating like forms with like meanings, in other words, of developing polysemies (Haiman 1985a). Since grammaticalization comprises the set of changes whereby structural relationships and associations among them are given grammatical expression, it is not surprising that it typically involves polysemy.

4.2.3 *Conversational and conventional inferencing*

With regard to pragmatics, we are particularly interested in those inferences that are made in linguistic contexts from one clause or constituent to another, or even from one utterance to another. These are in principle implicational inferences (in the linguistic jargon called “implicatures”) of the type characterized by Grice (1975) as “conversational.” Grice suggested that such inferences are computable on the basis not of lexical meanings alone, but of lexical meanings together with implicatures arising from speech act maxims such as “Make your contribution

as informative as is required (for the current purposes of the exchange)” (the first maxim of Quantity), “Do not make your contribution more informative than is required” (the second maxim of Quantity), “Try to make your contribution one that is true” (the maxim of Quality), “Be relevant” (the maxim of Relation), and “Be perspicuous” (the maxim of Manner) (Grice 1975: 45–7). There is considerable debate whether this is the right set of maxims. Indeed, it has been suggested that a principle of Relevance alone, defined in such a way as to include informativeness, is sufficient to account for pragmatic meaning (Sperber and Wilson 1995[1986]; for other views, see Atlas and Levinson 1981, Horn 1984, updated in Horn 1996 and Levinson 2000). As we will show, grammaticalization changes seem to draw primarily on the second maxim of Quantity, in the form “Say no more than you must and mean more thereby” (see Levinson 1983) and Relevance. Another point of debate has been whether “maxim” is the best term for the reasoning processes that people bring to the speech event, since it evokes imperatives and principles. We prefer the term “heuristic” proposed by Levinson (2000), which evokes strategic choice in the communicative situation.

Most conversational implicatures are strictly speaking interpreted abductively (given an utterance, hearers may relate it to a general heuristic, and guess the speaker’s intent). The guess may be wrong because the heuristics can always be flouted, e.g., it is possible for speakers to be uninformative or to lie. Furthermore, implicatures are “cancelable” either by the speaker (in which case an explanation is given), or by hearers’ inferences from the situation. An example is:

- (3) a. John has three cows.

The implicature from the first heuristic of Quantity is that:

- (3) b. John has three cows and no more.

But this can be canceled by a vague statement such as

- (3) c. John has three cows, if not more.

and even denied:

- (3) d. John has three cows, in fact ten.

The pragmatic effect of conversational implicatures across utterances and their cancelability can be illustrated by the example of conjoined clauses without any connective, that is, without any structural marker of coherence. These are likely to be interpretively enriched as having some coherence, that is, relevance to each other, simply because they are uttered in sequence.

- (4) a. The earthquake hit at 8 a.m. A four-car crash occurred.

Typically the relationship inferred will be that of temporal sequence and even causal connection if the clauses are action/event clauses and connectable in terms of encyclopedic or world knowledge, as in (4a); however, no such relationship is likely to be inferred if the sequence is incoherent in terms of world knowledge, as in (4b), where inference of a causal connection is unlikely, and even close temporal connection may be in doubt:

- (4) b. A four-car crash occurred. The earthquake hit at 8 a.m.

If a grammatical form is present, e.g., *and*, *because*, *you see*, this element will further “constrain the relevance of the proposition it introduces” (Blakemore 1987: 130). But this constraint still leaves indeterminacies. For example, the presence of *and* in (5a) implies only that a connection is intended by the speaker and that the hearer should compute one; it does not require that the implicature is a causal one, though that is what hearers would typically assume unless causality is canceled, as it is in (5b):

- (5) a. The earthquake hit at 8 a.m. and a four-car crash occurred.
 b. The earthquake hit and a four-car crash occurred, but actually the cause was the fog, not the earthquake.

Conversational implicatures are typically contrasted with “conventional” ones (in the next section we will show that conventional implicatures typically arise out of conversational ones). Conventional implicatures are unpredictable and arbitrary, that is, they must be learned as part of the polysemies of the word, and are not cancelable. Thus in (6a) the verb *manage* conventionally implicates (6b):

- (6) a. John managed to solve the problem.
 b. John solved the problem.

Neither (6a) nor (6b) are cancelable by, e.g.:

- (6) c. but he didn’t solve it.

Since they must be specially learned along with the phonological and syntactic characteristics for the item in question, conventional meanings can, at least for our purposes, be included among the semantic polysemies of a form. For a classic case consider the temporal and causal meanings of *since*, as in (7):

- (7) a. I have done quite a bit of writing since we last got together. (temporal)
 b. Since I have a final exam tomorrow, I won’t be able to go out tonight. (causal)

With *since*, when both clauses refer to events, especially events in the past, the reading is typically temporal, as in (7a). When one clause refers to a non-past

event or to a state, the reading is typically causal, as in (7b). The causal meaning is conventional and not cancelable, as illustrated by (7c):

- (7) c. *Since I have a final exam tomorrow, I won't be able to go out tonight, but not because of the exam!

In (7a) the first clause contains a past tense. In (7b) neither clause does. In other words, different meanings of *since* can be associated with different structural contexts. The difference between these meanings is sometimes syntactically obscured, and then there can be ambiguity, as in (7d):

- (7) d. Since Susan left him, John has been very miserable. (temporal or causal)

These facts allow us to conclude that *since* is semantically ambiguous (polysemous).

We turn now to the question of whether there are pragmatic as well as semantic polysemies. Consider, for example, *after* in (8a):

- (8) a. After we read your novel we felt greatly inspired.

This may be interpreted as a literal statement of temporal sequence, or it may implicate:

- (8) b. Because we read your novel we felt greatly inspired.

The implicature in (8b) strengthens informativeness because it enriches the relation between *After we read your novel* and the rest of the utterance, thus providing an interpretation of why the speaker thought it was relevant to include these temporal facts. However, there are no regular structural correlates for this relationship, such as completed versus ongoing eventhood, and there are no regularly associated, uncancelable causal implicatures. Rather, example (8a) suggests that there can be pragmatic ambiguities/polysemies as well as semantic ambiguities/polysemies. For fuller discussion of the importance of recognizing both pragmatic and semantic ambiguities, see Horn (1989) and Sweetser (1990).

4.3 The role of pragmatic inferencing in grammaticalization

With regard to the question of what role pragmatic inferencing has in grammaticalization, toward the end of his seminal article "Logic and conversation," Grice tentatively stated: "it may not be impossible for what starts life, so to speak, as a conversational implicature to become conventionalized" (1975: 58). This idea had been explored very briefly in Geis and Zwicky (1971) in connection with the development of causal *since*, and was explicitly followed up in Cole

(1975) in connection with *let's*. Dahl hypothesizes that many of what he calls “the secondary meanings” of tense and aspect, e.g., the “present relevance” of the perfect, are derived by conventionalization of implicatures. Below we will show that in early stages of grammaticalization conversational implicatures frequently become “semanticized,”² that is, become part of the semantic polysemies of a form. Dahl’s characterization of the process is as follows:

if some condition happens to be fulfilled frequently when a certain category is used, a stronger association may develop between the condition and the category in such a way that the condition comes to be understood as an integral part of the meaning of the category. (Dahl 1985: 11)

These insights have been elaborated in connection with semantic change in general, not specifically grammaticalization, in Traugott (1999) and Traugott and Dasher (2002). Here it must suffice to note that the kinds of semantic change undergone by lexical items or constructions in grammaticalization are a subset of semantic changes in general, and that for inferences to play a significant role in grammaticalization, they must be frequently occurring, since only standard inferences can plausibly be assumed to have a lasting impact on the meaning of an expression or to function cross-linguistically.

Among stereotypical inferences we may include the inferences of causality from temporal sequence that we have already discussed in connection with *since* and *after*, the well-known logical fallacy (abduction) characterized as *post hoc ergo propter hoc* ‘after this, therefore because of this.’ By contrast, we would not expect grammaticalization of such strictly local and idiosyncratic, highly contextualized, inferences as are exemplified by:

- (9) a. What on earth has happened to the roast beef?
 b. The dog is looking very happy. (understood to implicate that perhaps the dog has eaten the roast beef) (Levinson 1983: 126)

One question is when we can recognize conventionalization to have occurred. A brief look, before we proceed, at some early examples of the contexts in which the change of *since* (originally *sipþan*) came about will be useful in serving as a methodological caution, and show that it is essential to look beyond individual sentences to larger contexts before reaching too hasty conclusions that change has occurred. In OE texts before AC 1050 *sipþan* as a preposition was used almost exclusively to mean ‘from the time that, after.’ The standard causal was *for þæm þe* ‘for that that,’ originally a deictic expression. As a connective *sipþan* meant ‘from the time that,’ that is, it marked the lower temporal boundary of the event in the main clause, and signaled an overlap with some point in an earlier event.

In certain contexts, however, the modern reader may detect a causal implicature. For example, Mitchell (1985 (2): 352) cites (10) as a putative example of causal *sipþan*:

- (10) þa sipþan he irre wæs & gewundod, he ofslog micel
 then, after/since he angry was and wounded, he slaughtered much
 þæs folces.
 of-that troop (c. 880, Orosius 4 l.156.11)

When we read this sentence out of its larger context, the participle *gewundod*, being perfective and involving change of state, seems to favor a temporal reading, but the adjective *irre* favors a causal one, since it appears to express state. Nevertheless, a temporal reading with the adjective is plausible if we assume that *was* is inceptive–resultative or perfective, i.e., ‘had come to be’ rather than ‘was in a state.’ In other words, the adjective in this context can be interpreted as expressing a contingent rather than a general state (the latter would more probably have been expressed by *bið*). Indeed we find that this sentence occurs in the context of a narrative concerning the legendary Pyrrhic victory. In the battle a Roman soldier wounds an elephant in the navel; this elephant, having become enraged and wounded, wreaks mayhem on the army. *Sipþan* in (10), then, is best interpreted as a temporal (indeed it partially translates the Latin adverbial *postquam* ‘afterward’) allowing secondary causative implicatures. It is not a conclusive example of the conventionalizing of causative inference in Old English.

A more convincing example is:

- (11) Ac ic þe wille nu giet getæcan þone weg sipþan ðu ongistst þurh
 but I thee will now still teach that way since thou seest through
 mine lare hwæt sio soðe gesælð bið, & hwær hio bið.
 my teaching what that true happiness is, and where it is
 ‘But still I will now teach you the way since you see that true happiness comes
 through my teaching, and where it is.’ (c. 880, Boethius 36 104.26)

Here *sipþan* translates the Latin causative *quoniam* ‘because.’ But even without the Latin original we can assume it is causative since the context is non-narrative: the stative perception-mental verb ‘see, understand’ introduces an aspectual generic clause signaled by the verb *bið* instead of the contingency verb *is*. Although the causal inference is detectable in Old English in examples such as (11), so many other examples are undecidable that we cannot establish that the causal inference had truly become conventionalized at this period. The change appears to have occurred in the fifteenth century, when the form is attested frequently in stative and other non-completive environments where the temporal reading is blocked, as in (11).

A second question is what types of inferences are most characteristic of these early stages. Two different though mutually intersecting kinds, metaphorical processes and metonymic processes, have been much discussed in the literature, and it is to these that we now turn.

4.3.1 Metaphorical processes

Metaphorical innovation is one of the most widely recognized processes in meaning change. Standard examples of long-established metaphors include such utterances as (12) and (13):

- (12) Sally is a block of ice. (Searle 1993 [1979]: 97)
 (13) The sentence was filled with emotion. (Reddy 1993 [1979]: 288)

Although definitions of metaphor vary, most have certain concepts in common, especially understanding and experiencing one kind of thing in terms of another, and directionality of transfer from a basic, usually concrete, meaning to one more abstract (see, for example, J. D. Sapir 1977; G. Lakoff and Johnson 1980; Claudi and Heine 1986; Heine, Claudi, and Hünemeyer 1991a; Kövecses 2002). Metaphorical processes are processes of inference across conceptual boundaries, and are typically referred to in terms of “mappings,” or “associative leaps,” from one domain to another. The mapping is not random, but motivated by analogy and iconic relationships.

These relationships tend to be observable cross-linguistically. Some have been thought of as “image schemata” with very concrete sources that are mapped onto abstract concepts. In the lexical domain one image schema that is well known is that of seeing and knowing, grasping and understanding (as in *I see/grasp the point of your argument*). In this particular case the relatively concrete concept has been said to be mapped onto the relatively abstract one in a relationship called the “mind-as-body metaphor”: bodily experience is a source of vocabulary for psychological states (Sweetser 1990: 28–48). Another well-known relationship is “force dynamics” (Talmy 1988; Jackendoff 1990; Sweetser 1990), the relationship of forces and barriers found in such expressions as (14) and (15):

- (14) The crack in the stone let the water flow through.
 (15) I have a mental block about sports.

Metaphoric processes have traditionally been regarded as semantic. Recently, however, it has been suggested that, not being truth conditional, but rather being based in communicative use, they are more appropriately considered pragmatic (e.g., Levinson 1983). We accept the view that metaphor is pragmatic, and argue below that in so far as metaphor is primarily analogical in character, it is different

from the kinds of conversational processes based on heuristics mentioned above, which operate primarily in linear, syntagmatic ways. A further difference identified by Green (1996 [1989]: 122) is that, at a superficial level, metaphors often involve propositions that are intended to be recognized as literally false (for example, it is false that a person can be a block of ice), but conversational implicatures do not.

Most examples of metaphorical processes in language change have been discussed with respect to the lexicon. However, arguments have been put forward that early grammaticalization is also strongly motivated by metaphoric processes. Typical of early claims along these lines is: "Rather than subscribe to the idea that grammatical evolution is driven by communicative necessity, we suggest that human language users have a natural propensity for making metaphorical extensions that lead to the increased use of certain items" (Bybee and Pagliuca 1985: 75).

Probably the most appealing examples of metaphoric processes in grammaticalization are provided by the development of spatiotemporal terms. Claudi and Heine (1986) and Heine, Claudi, and Hünemeyer (1991a,b) discuss the development of body part terms into locatives, of spatial terms into temporals, etc. in terms of metaphors such as *SPACE IS AN OBJECT*, *TIME IS SPACE* (capitals indicate abstract, cross-linguistic meanings, as opposed to language-specific lexical items). For example, spatial terms such as *BEHIND* can be derived metaphorically from a body part (an example of the shift from *OBJECT* > *SPACE*), and subsequently temporal terms can be derived metaphorically from the spatial term (via *SPACE* > *TIME*), e.g., *We are behind in paying our bills*. Spatial terms abound cross-linguistically as temporal particles, auxiliaries, etc. (see Traugott 1978, 1985a; Bybee and Dahl 1989; Bybee, Pagliuca, and Perkins 1991 on expressions of the future). A few examples from English which have been regarded as metaphorical in origin include *be going to* (future), *in the years ahead* (future), *drink something up/down* (completive), *drink on* (continuative), *come to believe that* (ingressive). Extensive examples from African languages can be found in Heine, Claudi, and Hünemeyer (1991a,b), and from Oceanic languages in Lichtenberk (1991b) (e.g., *GO* for continuative and future, *COME* for ingressive and future). Svorou (1993) and Haspelmath (1997) provide detailed cross-linguistic evidence for *SPACE* > *TIME*.

Another domain of meaning change among grammatical categories that has been widely regarded as metaphoric is the development of modal meanings, particularly the development of meanings relating to obligation into meanings relating to possibility and probability (known as "epistemic" meanings). For example: "The obligation sense of *have to* predicates certain conditions on a willful agent: X is obliged to Y. The epistemic sense is a metaphorical extension of obligation to

apply to the truth of a proposition: X (a proposition) is obliged to be true” (Bybee and Pagliuca 1985: 73). In this view (modified in Bybee 1990), the process envisioned appears to be strictly speaking the schematic mapping of one concept onto another. Building on Talmy (1976, 1988), Sweetser takes a different approach to the modals, that of “sociophysical concepts of forces and barriers” (1990: 52). The *may* of permission is, according to Sweetser, understood in terms of “a potential but absent barrier,” obligative *must* in terms of “a compelling force directing the subject towards an act.” The force of *must* is “directly applied and irresistible,” whereas that of *have to* is resistible under certain circumstances, cf.:

- (16) I have to/??must get this paper in, but I guess I’ll go to the movies instead.
(Sweetser 1990: 54)

Sweetser regards the epistemic meanings of these modals as deriving from the tendency to experience the physical, social, and epistemic worlds in partially similar ways. This similarity in experience, she suggests, allows the mapping of sociophysical potentiality onto the world of reasoning. For example, with respect to *may*, she says: “In both the sociophysical and the epistemic world, *nothing prevents* the occurrence of whatever is modally marked with *may*; the chain of events is not obstructed” (p. 60). With respect to *must*, she gives the following analysis (p. 61):

- (17) a. You must come home by ten. (Mom said so.)
‘The direct force (of Mom’s authority) compels you to come home by ten.’
b. You must have been home last night.
‘The available (direct) evidence compels me to the conclusion that you were home.’

She goes on to show that yet another metaphorical mapping is possible: of potential barriers to the conversational world (what is often called metalinguistic expression). Thus, in a hypothetical situation where Mondale’s advisor is giving directions to a speech writer, the following might be imagined (p. 71):

- (18) Reagan will/must be a nice guy (as far as the content of the speech is concerned, even if we criticize his policies).

In other words, “the interlocutor is being allowed to treat a certain statement as appropriate or reasonable.”

Can all possible types of metaphor be drawn on in grammaticalization? The answer appears to be no. Talmy (1983, 1988, 2000) has suggested that only certain types of spatial concepts are used cross-linguistically in grammatical items: specifically, topological concepts. Thus precise distances between points on a scale, or precise angles, do not grammaticalize. Indeed, angles in general (e.g., *corner in*

time) typically do not appear to grammaticalize.³ However, topological relations on a linear parameter frequently do so, e.g., *front-back*, *up-down*.

Sweetser has suggested that when a lexical item expressing a spatial concept is grammaticalized, only the topological concept is transferred. The concrete image associated with the lexical item is replaced by a more schematic one, and the meaning transfer “is to a fairly abstract, topological domain... so there is less fleshing-out of meaning” (Sweetser 1988: 393).

4.3.2 Metonymic processes

There is little doubt that metaphor is one process at work in grammaticalization. However, since reanalysis, not analogy, has for long been recognized as the major process in grammaticalization at the structural, morphosyntactic level, it would be surprising if metaphor, which is analogical, were the prime process at work pragmatically and semantically. In this section we show that other processes, which depend on contiguity and association in the flow of speech, also play a major part, and that some instances of grammaticalization that have heretofore been regarded as metaphorical can be seen to arise out of semantic contiguity rather than or as well as out of semantic analogy.

The overriding importance that metaphor was given in many discussions of grammaticalization during the 1980s and early 1990s seems to have derived in part from the tendency to think in terms of “lexical item > grammatical item,” i.e., in terms of form, relatively independently of context rather than in terms of “use of lexical item in discourse > grammatical item,” i.e., in terms of form in utterance contexts.⁴ For example, when the lexical item *go* is considered out of context and is said to grammaticalize to an auxiliary, metaphor is naturally invoked with respect to its spatial properties. But in fact it was not *go* that grammaticalized; the phrase *be going to* did, presumably only in very local contexts, e.g., that of *be going in order to V*. The contiguity with *to* in the purposive sense must have been a major factor in the development of the future meaning in *be going to* as an auxiliary (we discuss this point more fully below).

Another reason for the focus on metaphor was presumably also that the term metonymy had until that time been thought to be rather insignificant. For example, Dirven speaks of metaphor as a “major associative leap” but of metonymy as a “minor process” (1985: 98). Furthermore, in the tradition deriving from Jakobson and Halle’s (1956) classic distinction between metaphor as choice functioning on the paradigmatic axis versus metonymy as association and sequence functioning primarily on the syntagmatic axis, metaphor was thought to lead to homogeneity and coherence, metonymy to juxtaposition and potential incoherence (J. D. Sapir 1977: 4). The term was used primarily for changes arising out of contiguity in

the non-linguistic world, cf. such examples as Lat. *coxa* ‘hip’ > Fr. *cuisse* ‘thigh’ (the parts of the body are spatially contiguous in the physical world), and *boor* ‘farmer’ > ‘crude person’ (association of behavior with a certain person or class of persons). One of the most famous examples is the transfer by association of the term for ‘prayer’ (OE *gebed*) to the objects by which a series of prayers was counted, the *beads* of a rosary (and ultimately, by generalization, to any beads). However, contiguity in the utterance, often resulting in ellipsis, had also been used as an example of metonymy, cf. French *foie* ‘liver’ < Latin *iecur ficatum* ‘liver fig-stuffed.’ Neither of these senses of metonymy is useful for the study of grammaticalization. However, conceptual association also involves metonymy, and in this sense (often known as “conceptual metonymy”) the term is valuable. This sense can already be found in Stern, who, in speaking of “permutation,” says it results from “a word [being] used in a phrase where a notion in some way connected with its meaning is liable to form an element of the context” (1931: 353). He goes on to list under examples of permutation the development of the logical meanings of *considering*, *supposing*, and of concessive *while*. More recently Anttila suggested that “[m]etaphor is semantic transfer through a similarity of sense perceptions,” and is analogical and iconic, while metonymy is semantic transfer through contiguity and “indexical” (1989 [1972]: 141–2). In one of the first works in the last few decades to recognize the importance of metonymy in grammatical change, Brinton (1988) argued that the development of the English aspect markers, including *have*, is metonymically rather than metaphorically motivated.

Recently the fundamental importance of conceptual metonymy in language in general has been widely recognized. Indeed, there has been a major shift in thinking and it is coming to be increasingly recognized as “probably even more basic [than metaphor] to cognition” (Barcelona 2000: 4). As a cognitive process in which “one conceptual entity . . . provides access to another conceptual entity . . . within the same domain’ (Kövecses and Radden 1998: 38), metonymy points to (“indexes”) relations in contexts that include interdependent (morpho)syntactic constituents. In an utterance such as (19) the verb *go* invites the conversational inference that the subject arrived at a later time at the destination, and the purposive *to*, introducing a subordinate clause, invites the conversational inference that someone intended the marriage to occur:

- (19) I was/am going to be married. (in the sense ‘I was/am going for the purpose of getting married’)

However, this implicature can be canceled:

- (20) I was going/on my way to be married, but on the plane I changed my mind and decided to join the Army.

We hypothesize that the future meaning of *be going to* was derived by the semanticization of the dual inferences of later time indexed by *go* and purposive *to*, not from *go* alone. Indeed, we hypothesize that the inference from purposive *to* must have played a significant role in the grammaticalization of *be going to* given that the major syntactic change involved in the development of the auxiliary is the rebracketing of [[... be going] [to S]] as [... be going to V X] (Section 1.1). The progressive *be -ing* indexed activity in process, and so motivated the tendency for *be going to* to be interpreted as a purposive that was relevant to the reference time of the clause and likely to be imminent (see Bybee and Pagliuca 1987; Pérez 1990, who differ from the analysis presented here mainly in treating the change as a case of metaphorization).

To appreciate the importance of the relationship between *to* and *go*, in the development of auxiliary *be going to*, consider the following possible early instance:

- (21) Thys onhappy sowle... was goyng to be broughte into helle for the synne and onleful [unlawful] lustys of her body.
(1482, Monk of Evesham [*OED* go 47b])

This can be understood as an expression of motion in the context of the belief that after death the soul goes on a journey with the purpose of being rewarded or punished for actions in life. Note that in this example the passive demotes the inference that the subject of *go* is volitional or responsible with respect to the purposive clause. Because the destination of the journey (hell) is an adjunct not of *goyng to* but of *broughte*, the directionality of *going* is also demoted, and the inference of imminent future resulting from the purposes of the judges of the dead is promoted.

Similarly, in the passage in (22) the answer to *whither away* is (*to*) a messenger, and *I am going to deliver them* seems best understood as answering the question (*why*) *so fast?*, in other words, it seems more informative if it is inferred to answer the question in terms of purposes rather than directions:

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

The full semanticization (and grammaticalization) of *be going to* is evidenced when the following subject and/or the verb is incompatible with purposiveness, for example, an inanimate subject or a verb of mental experience such as *hear*, or *like*. Once the semanticization of later time/future had occurred, the *will* future

could no longer be used with *be going to*, presumably because it had become partially redundant, and did not fit the auxiliary verb structure into which the construction had been absorbed. (Note, however, that the *will*-future can still occur in the main verb construction *be going to*, as in *I will be going to visit Aunt Mildred tomorrow*.)

The metaphor account, whereby a trajectory through space is mapped onto a trajectory in time, does not give adequate insight into why the progressive and most especially *to* are involved in the English expression *be going to*. This becomes particularly clear when we compare the cross-linguistic grammaticalization of the verb with the abstract meaning GO to future markers, each of which seems to have a slightly different history. Sometimes, for example, there is no overt purposive, in which case the future inference arises out of the directional verb and its associated aspect alone, as in French. Sometimes, however, GO may be grammaticalized into either a purposive or a temporal marker of imminence, as occurred to **bang* ‘go’ in Rama:

- (23) a. Tiiskama ni-sung-bang taak-i.
 baby I-see-SUB go-TNS
 ‘I am going in order to see the baby.’
 b. Tiiskama ni-sung-bang.
 baby I-look-at-ASP
 ‘I am going to look at the baby.’ (Craig 1991: 457)

In each case inferences from the highly local contexts of the verb in its linear position within the clause appear to be the immediately motivating factors for change, though the capacity to create metaphors of time from space may well provide a cognitive framework that supports the changes.

We have suggested that semanticization of conversational inferences played a major role in the development of *be going to*. Another example is provided by the development of *while* (see Traugott and König 1991, on which the following discussion is based). This connective originated in OE in an adverbial phrase translatable as ‘at the time that’ consisting of the accusative distal demonstrative, the accusative noun *hwile* ‘time,’ and the invariant subordinator *þe*, a highly explicit coding of simultaneity, as in:

- (24) & wicode þær þa hwile þe man þa burg worhte
 and lived there that:DAT time:DAT that one that fortress worked-on
 & getimbrode.
 and built
 ‘And camped there at the time that/while the fortress was worked on and built.’
 (Chron A [Plummer] 913.3)

This phrasal expression was reduced by late OE to the simple conjunction *wile*:

- (25) Dæt lastede þa [xix] winttre wile Stephne was king.
 ‘That lasted those 19 winters while Stephen was king.’
 (ChronE [Plummer] 1137.36)

In the process, the precise specification of simultaneity signaled by the demonstrative was lost, allowing for other, less precise, conversational inferences to play a part. One such inference is that the conditions specified in the subordinate clause serve not only as the temporal frame of reference for those in the main clause, but also as the grounds for the situation (the disasters lasted nineteen years because Stephen was king). Such an inference to grounds for the situation is dominant over temporality in some examples dating from the later fourteenth century:

- (26) Thar mycht succed na female,
 Quhill foundyn mycht be ony male.
 ‘No female was able to succeed while any male could be found.’
 (1375, Barbour's Bruce 1.60 [OED **while** 2a])

The causal inference from *while* did not become semanticized in English. However, in some languages this inference to the grounds for the situation has become the main extension of **WHILE**. For example, in German the temporal meaning of *weil* ‘during’ has become obsolete and the causal has become the main meaning; interestingly, causal *weil* came to be used in the latter part of the twentieth century with a concessive meaning (main clause rather than subordinate clauses syntax) (Günther 1996).

In English a different inference came to dominate, that of surprise concerning the overlap in time or the relations between event and ground. This led to the adversative, concessive meaning (cf. similar developments for *as long as*, *at the same time as*). Probable instances of the semanticization of surprise and hence concessivity appear in the early seventeenth century, among them:

- (27) Whill others aime at greatnes boght with blod,
 Not to bee great thou stryves, bot to bee good.
 ‘While others aim at greatness that is bought with blood, you strive to be not great but good.’ (1617, Sir W. Mure, *Misc. Poems* xxi.23 [OED **while** 2b])

This could be interpreted as a statement about simultaneous behaviors. However, there is a strong inference, reinforced by the inversion in the second line, that it is unusual not to be bloodthirsty. Unambiguous examples of concessive meanings appear later, typically with present-tense stative verbs, e.g., *While you like peaches, I like nectarines*. The overall shift of *while* is from reference to a relatively concrete state of affairs (a particular time) to expression of the speaker’s assessment of the

relevance of simultaneity in describing events, to assessment of contrast between propositions. In other words, it demonstrates a shift to a relatively abstract and subjective construal of the world in terms of language.

A similar development to increased subjectivity is evidenced by *be going to*; the motion verb requires that the direction of motion be anchored in the subject as well as in the speaker's viewpoint.⁵ But the auxiliary can be anchored in the speaker's subjective viewpoint alone, not in that of the subject:

- (28) An earthquake is going to destroy that town. (Langacker 1990: 23)

An earlier example is:

- (29) It seems as if it were going to rain. (1890, Cham. Jnl. [*OED go* V.b])

Other examples of subjectification include the development of epistemic modals, for example the development of *must* in the sense of 'I conclude that' as in (17b) from *must* in the sense of 'ought' as in (17a) (Traugott 1989; Langacker 1990).

4.4 Metaphor and metonymy as problem solving

In discussing the principle of exploiting old means for novel functions, and the recruitment of concrete for more abstract terms, Heine, Claudi, and Hünemeyer suggest that:

grammaticalization can be interpreted as the result of a process which has **problem-solving** as its main goal, its primary function being conceptualization by expressing one thing in terms of another. This function is not confined to grammaticalization, it is the main characteristic of metaphor in general.

(1991b: 150–1)

In other words, semantic change in general, not just grammaticalization, can be interpreted as problem solving (see also C. Lehmann 1985). One problem to be solved is that of representing members of one semantic domain in terms of another, and metaphoric strategies serve this purpose. The second problem is the search for ways to regulate communication and negotiate speaker–hearer interaction. We have shown that this is a kind of metonymic change, indexing or pointing to meanings that might otherwise be only covert, but are a natural part of conversational practice. The main direction of both types of problem solving is toward informativeness, but the two types correlate with shifts along different axes. Metaphorical change involves specifying one, usually more complex, thing in terms of another not present in the context. Metonymic change,

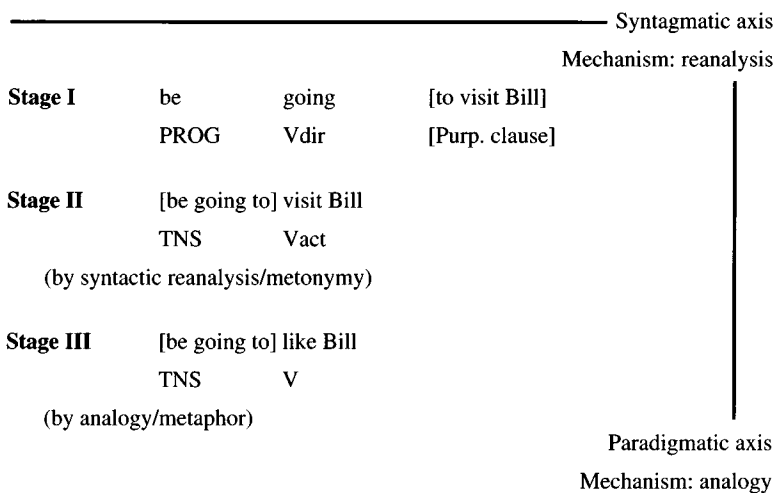


Figure 4.1 Revised schema of the development of auxiliary *be going to*

on the other hand, involves specifying one meaning in terms of another that is present, even if only covertly, in the context. It is largely correlated with shifts to meanings situated in the subjective belief state or attitude toward the situation, including the linguistic one. While metaphor is correlated primarily with solving the problem of representation, metonymy and semanticization of conversational meanings are correlated with solving the problem of expressing speaker attitudes.

In summary, metonymic and metaphorical inferencing are complementary, not mutually exclusive, processes at the pragmatic level that result from the dual mechanisms of reanalysis linked with the cognitive process of metonymy, and analogy linked with the cognitive process of metaphor. Being a widespread process, broad cross-domain metaphorical analogizing is one of the contexts within which grammaticalization operates, but many actual instances of grammaticalization show that conventionalizing of the conceptual metonymies that arise in the syntagmatic flow of speech is the prime motivation for reanalysis in the early stages. We can now refine the model for *be going to* outlined in Figure 3.2. to specify that syntagmatic reanalysis is accompanied by metonymic strategies, and paradigmatic, analogical change by metaphorical ones. The revised model is presented in Figure 4.1.

The competing motivations of expressivity (which underlies metonymic and metaphorical inferencing) and routinization, together with the mechanisms of reanalysis and analogy discussed in Chapter 3, will be shown in the next chapter to motivate the unidirectionality typical of grammaticalization.

4.5 Pragmatic enrichment versus “bleaching”

From very early times researchers on issues related to grammaticalization have observed that it involves loss of semantic content. This has been described by the metaphor of “fading” or “bleaching” (Gabelentz spoke of “verbleichen” ‘to grow pale,’ Meillet of “affaiblissement” ‘weakening’). More recently, Heine and Reh characterized grammaticalization as: “an evolution whereby linguistic units lose in semantic complexity, pragmatic significance, syntactic freedom, and phonetic substance” (Heine and Reh 1984: 15). Readers will have noted that in this chapter we have, however, spoken of pragmatic enrichment, strengthening, and so forth. This is because we have been discussing the beginnings of grammaticalization, that is, the motivations that permit the process to begin, rather than its outcomes. There is no doubt that, over time, meanings tend to become weakened during the process of grammaticalization. Nevertheless, all the evidence for early stages is that initially there is a redistribution or shift, not a loss, of meaning.

For example, with reference to the development of future *go*, Sweetser says: “we lose the sense of physical motion (together with all its likely background inferences). We gain, however, a new meaning of future prediction or intention – together with *its* likely background inferences” (Sweetser 1988: 392). In speaking of the subjectification of *be going to*, Langacker draws attention to the loss of objective locational reference points that movement entails, and suggests that this loss is replaced by realignment to the speaker’s temporal perspective (1990: 23). In other words, one meaning is demoted, another promoted.

As grammaticalized forms become increasingly syntacticized or morphologized they unquestionably cease over time to carry significant semantic or pragmatic meaning. This can most clearly be seen when former lexical items become empty syntactic elements, as in the case of *do*, or when formerly separate morphemes become bound and serve primarily as “morphological detritus” after repeated fusion (see Chapter 6). An excellent example is provided by the development of French *ça* ‘that,’ a form which is the worn-down relic of several stages of expressive reinforcement:

- (30) hoc ‘that’ > (ecce) hoc ‘behold that’ > eccehoc > ço > ce > ce(là) ‘that there’ > celà > ça (Lüdtke 1980: 212)

The individual meanings of *hoc*, *ecce*, and *là* have been lost, as has the form’s distal demonstrative function (M. Harris 1978: Chapter 4).

Two general working principles arise out of our understanding of the processes of inferencing in grammaticalization. One is that the meanings will always be derivable from the original lexical meaning by either metaphorical or conceptual metonymic inferencing. Therefore meaning changes in grammaticalization are not

arbitrary. Secondly, since the initial phase of grammaticalization involves a shift in meaning, but not loss of meaning, it is unlikely that any instance of grammaticalization will involve a sudden loss of meaning.

One of the most well-known examples of grammaticalization in English, the development of “empty” *do*, has been treated mainly as an example of syntactic change, and might be regarded as a counterexample. There has been considerable speculation about its origins. Causative *do*, as in (31), has been thought to be one source (Ellegård 1953):

- (31) þe king sende efter him & did him gyuen up ðat abbotrice of Burch.
 ‘The king sent for him and made him give up the abbey of (Peter)borough.’
 (Peterb. Chron 1132 [*MED don* 4a])

Denison (1985) and Stein (1990a) have suggested that grammaticalization occurred not simply via loss of causative meaning, but via a perfective meaning, which can arise via pragmatic strengthening particularly in past tense causative contexts. For example, we can interpret *dide him gyuen up ðat abbotrice* in (31) as representing not only that the king’s making him do something occurred in the past, but also that the action (give up his abbey) was achieved. This perfective meaning demotes the issue of who undertook the action, and permits the inference that the action is the significant issue. This is particularly likely to occur in constructions without an overt subject of the non-finite clause following *do*. For example, in (32) the subject (‘they’) may have caused the women to be mocked, but the absence of an overtly expressed intermediary party between the mockers and their victims demotes the causality and promotes the accomplishment (perfectivity) of the action:

- (32) And so thei dede bothe deseiu ladies and gentilwomen, and bere forthe diuerse langages on hem.
 ‘And so they both mocked ladies and gentlewomen, and made various allegations against them.’
 (c. 1450, Knt. Tour-L, 2.24; cited in Denison 1985: 50)

There are some problems with the causative analysis, however, among them the fact that many instances of *do* in Middle English in the area in which it developed (southwest England) were and continue to be habitual in meaning (Garrett 1998). Garrett cites several modern examples such as this one from Somerset:

- (33) The surplus milk they did make into cheese and then the cheese did go to the different markets, that’s how that did work.
 (Garrett 1998: 292, citing Ihalainen 1976: 615)

Whatever the final analysis turns out to be, the point is that sudden emptying of meaning is not expected. Furthermore, it is incumbent on the researcher to seek a plausible set of inferences that enable changes to occur.

Perhaps the most damaging evidence against the automatic association of bleaching and sudden emptying of meaning with grammaticalization comes from evidence that later constraints on structure or meaning can only be understood in the light of earlier meanings. In other words, when a form undergoes grammaticalization from a lexical to a grammatical item, some traces of its original lexical meanings tend to adhere to it, and details of its lexical history may be reflected in constraints on its grammatical distribution. This phenomenon has been called “persistence” (Hopper 1991).

An example is provided by Lord’s discussion of the development in West African languages of object markers (“accusative cases”) out of former serialized verbs like ‘take.’ In Gã (a Benue-Kwa language of West Africa), the form *kè* is an accusative case marker in sentences such as:

- (34) È kè wòlò ñmè-sì.
 she ACC book lay-down
 ‘She put down a book.’ (Lord 1993: 118)

It was originally a verb meaning ‘take’ (see Lord 1993: 53–6) which has become grammaticalized as an accusative case marker, and sentences such as these are historically of the type ‘He took a book [and laid [it] down.’ Certain restrictions on the distribution of this case marker can only be understood from the point of view of its origin in the meaning ‘take.’ Let us for a moment consider a language which has a fully developed accusative case marker, Latin. In Latin, accusative is a general marker of direct objects. The Latin accusative case occurs (to a large extent) irrespective of the semantic relationship between the verb and its object. Accusative objects in Latin may be perceived, produced, ordered, imagined, and so on. With a few exceptions, any noun which has the role of an object is marked as an accusative.

Gã is quite different from Latin in this respect. Consider, for example, (35a) and its ungrammatical counterpart (35b):

- (35) a. È ñmè wòlò.
 she lay egg
 ‘She laid an egg.’
 b. *È kè wòlò ñmè.
 *she ACC egg lay (Lord 1993: 120)

The semantic relationship between verb and object in ‘laid down the book’ is quite different from that in ‘laid an egg.’ In the first, the object is changed (moved, grasped, etc.) through the action of the verb; the object is “affected” (Lord 1982; Hopper 1986a). In the second, the object is produced or brought about by the action of the verb; the object is “effected.” The accusative case marker in Gã can only be used if the object is “affected.” For this same reason, the accusative case marker *kè* is inappropriate if the verb is one of experiencing:

- (36) a. Tètè nà Kòkǎ.
 Tete saw Koko
 ‘Tete saw Koko.’
 b. *Tètè kè Kòkǎ nà.
 *Tete ACC Koko saw (Lord 1993: 120)

These kinds of constraints exist because Gã retains the restriction on accusative case marking which derives from the historical antecedent of this grammatical morpheme in the lexical verb ‘to take’: only objects which can be ‘taken’ are marked morphologically as accusatives.

Persistence of old meanings is a common phenomenon. Some differences in the meanings of the Present-Day English tense/modal auxiliaries reflect possibilities of meaning which have existed for over a thousand years. Bybee and Pagliuca show that: “the differences in the uses of these future markers [i.e., *will*, *shall*, *be going to*, PH and ET] can be understood as continuations of their original lexical meanings” (Bybee and Pagliuca 1987: 117). As mentioned in connection with example (2), there are several polysemies of the “future” *will*. These include prediction (the “pure” future), willingness, and intention. Bybee and Pagliuca show that two of these meanings were already implied in Old English. (37) exemplifies willingness, (38) intention:

- (37) Gif he us geunnan wile, þæt we hine swa godne gretan moton...
 if he us grant will, that we him so generous greet should...
 ‘If he will/is willing to grant that we should greet him who is so gracious...’
 (*Beowulf* 346–7;⁶ cited in Bybee and Pagliuca 1987: 113)
- (38) Wen’ ic þæt he wille, gif he wealdon mot, in þæm guðsele
 think I that he will, if he prevail may, in the war-hall
 Geotena leode etan unforhte.
 of-Geats men eat unafraid
 ‘I believe that he will, if he should prevail, devour the people of the Geats
 without fear in their war-hall.’
 (*Beowulf* 442–4; cited in Bybee and Pagliuca 1987: 113)

The “predictive” future has developed out of the intention/promise use of *will*. Bybee and Pagliuca show that the future meaning becomes established in the Middle English period when inanimates incapable of volition begin to appear as the subjects of *will* (see also Aijmer 1985). When this happened, it did not result in an across-the-board re-semanticization of *will*; the predictive future remains only one of several distinct meanings of *will* in Present-Day English. All that happened was that a new meaning was added to an already polysemous form, and thus new distributional possibilities were opened up for the form. Furthermore, the fact that *will* and *be going to* do not mean the same thing (*be going to* expresses present orientation and a goal-directed plan), is attributable to the fact that the latter had progressive and directional origins while the former did not.

The process of demotion of some lexical meanings and promotion of others is characteristic of semantic change in general. Those lexical meanings that are promoted in grammaticalization tend to be relatively abstract, and particularly relevant to expression of temporality, role relationships, connectivity, etc. (i.e., “grammatical meanings”). Furthermore, they tend to be the ones most salient in the original contexts/formulae within which grammaticalization takes place (e.g., intention regarding the future is salient in the contexts relevant to the grammaticalization of *be going to*). These originally salient meanings tend to persist over time and to constrain the later uses of the grammaticalized form. “Bleaching” must therefore be taken to be a very relative notion, and one that pertains almost exclusively to late stages of grammaticalization. An important question for future research is what exactly constitutes bleaching, when it can be said to set in, and how it correlates with morphosyntactic generalization.

4.6 Conclusion

We have argued that grammaticalization can be thought of as the result of the continual negotiation of meaning that speakers and hearers engage in. The potential for grammaticalization lies in speakers attempting to be maximally informative, depending on the needs of the situation. Negotiating meaning may involve innovation, specifically, pragmatic, semantic, and ultimately grammatical enrichment. It is largely motivated by metonymic and metaphorical inferencing. These are complementary processes linked with the dual mechanisms discussed in Chapter 3: reanalysis (linked with metonymy), and analogy (linked with metaphor). As innovations come to be adopted by members of the community, they may be subject to maximization of signal simplicity, and ultimately to various types of reduction, typically semantic bleaching, morphological bonding, and phonological attrition. It is to the unidirectional changes in structuration that arise from these diachronic processes that we now turn.