Gradience, gradualness and grammaticalization: How do they intersect? 
Elizabeth Closs Traugott and Graeme Trousdale 
Stanford University and The University of Edinburgh

Abstract

This volume is intended to address three questions: 1) How are we to understand the intersection between synchronic gradience and grammaticalization? 2) What insights does grammaticalization offer for assessing the validity of Aarts’s (2007) claims regarding synchronic gradience, specifically that there is a significant distinction between subsective and intersective gradience? 3) What does the intersection between grammaticalization and synchronic gradience tell us about the hypothesis of structural gradualness, and about whether work on grammaticalization needs reanalysis and analogy/extension, or some other mechanism? In this paper we present an overview of what we consider to be central issues in answering these questions and in developing a theory of micro-changes.

1. Introduction

In this chapter, we are concerned with some issues in discussions of linguistic gradience, the extent to which diachronic change may be said to be gradual, and the relationship between a gradient system and gradualness in change, from the perspective of grammaticalization understood as the development of grammatical functions. Speciﬁcally, we are interested in how gradience in the synchronic system (Aarts 2004, 2007a, b; Croft 2007) intersects with the gradual changes which appear to be characteristic of grammaticalization, on the assumption that: “changes are always manifested in synchronic variation” (Andersen 2001: 228). We argue that while there are some overlaps between gradience and grammaticalization, (i) evidence from grammaticalization causes problems for accounts of gradience which focus solely on morphosyntactic distribution and (ii) the intersection of grammaticalization and gradience is only a small part of the emergence of grammatical constructions. Therefore we suggest that some current formulations of gradience might benefit from some rethinking, based on evidence from diachronic grammaticalization. Similarly, theories of grammaticalization might also beneﬁt from better understanding of the concepts of abruptness and gradualness. This is especially the case in light of the ongoing debate regarding the relative importance of reanalysis and analogy in grammaticalization (see inter alia, Hopper and Traugott 2003, Lehmann 2004, Fischer 2007, Roberts 2007, Kiparsky Forthcoming).

The term ‘gradience’ is used in a number of ways in linguistic discourse. We are primarily concerned with two ways in which the term has been used. One pertains to the nature of boundaries between categories, the other to the organisation of members within a category. Together, these may be characterised as the “(perceived) interlacing of the categories of the language system” (Aarts 2004: 5). For reasons of space we exclude a number of alternative conceptualizations of, and solutions to problems of, gradience in grammar. In particular, we do not address: (a) gradience as degrees of ‘grammaticality’ understood as the well-formedness of a particular utterance (Sorace and Keller 2005; see also the collection of articles in Fanselow, Fery, Schlesewsky and Vogel 2006); (b) gradience as frequency effects. The latter may be correlated with gradience as degrees of well-formedness or category membership (an issue sometimes addressed within OT and probabilistic approaches to grammar, e.g. Bod, Hay and Jannedy 2003), or with constraints on collocations (Hilpert 2008).

---

1 Many thanks to Hendrik De Smet for many insightful comments on an earlier draft of this paper, and to David Denison, Ian Roberts, and Anette Rosenbach for discussion of the issues.

2 A definition of grammaticalization as increase in morphosyntactic dependency (see Haspelmath 2004) might suggest different issues.
We are by no means the first to consider the relationship between gradience, gradualness and grammaticalization. DeLancey (1997) cites grammaticalization as a reason for the need to consider syntactic categories to be gradient. Lightfoot argues that:

Given that the abstract layers and sub-layers of language (i.e. lexicon, phonology, morphology, syntax, semantics, etc.) can be interpreted as gradiently related categories, we may claim that the dynamic processes which utilize said categories are likewise of a gradient nature. (Lightfoot 2005: 590)

Unlike Lightfoot, we do not consider diachronic processes themselves to be gradient. Rather, we argue that most instances of change involve small micro-steps that are in fact discrete and therefore abrupt (in a tiny way) (see §4). However, because different parts of a construction may undergo changes at different points in time, the change to the construction as a whole may appear to be gradual.3

In establishing how to model syntactic gradience, and in attempting to answer whether such gradience is the product of gradual diachronic change, we may contrast formal, generative approaches to change with functional accounts. The former typically see change as the product of abrupt syntactic reanalyses (Lightfoot 1979, 1999, Roberts 1993, 2007; Roberts and Roussou 2003). By contrast, functional approaches to change see the process as gradual, such that every “intermediate step in the process [of syntactic change: ET/GT] represents an intermediate construction type in structural terms” (Croft 2001: 313). Similarly, proponents of emergent grammar (Hopper 1987, Bybee and Hopper 2001) have also stressed that synchronic gradience is an effect of diachronic gradualness. Differences between generative and emergentist views on gradience and gradualness highlight a further central question: what is meant by ‘change’? For generative grammarians, what changes is (parameterized configurations of) grammars, and change happens during acquisition (see Kiparsky 1968, Lightfoot 1999). For most functionalists, structural changes are aligned with patterns of use, and usage shapes grammatical representation (Croft 2000; Bybee 2006). Such structural changes occur across the lifespan, and spread across social networks (Milroy 1992) or the speech community (Labov 2001).4 Some formalists exclude gradience not only from matters of linguistic form, but also from matters of function (Bouchard 1995: 33). While we recognise that the gradual spread of change is both ‘structural’ (i.e. through the linguistic system) and ‘social-contextual’ (i.e. transmitted across groups of speakers and texts, at different levels of formality, in different locations, at different times), we will focus here only on the former (see particularly §2.2). However, we acknowledge the importance of the latter (on which see further Bybee and Hopper 2001, Bybee 2003, 2006, 2007). The assumptions made in the rest of this chapter are broadly consistent with functional approaches to change in general and grammaticalization in particular.

We suggest programmatic answers to the three questions posed in the preface. First we provide some definitions of gradience (§2.1), and gradualness (§2.2). Regarding the first question about the intersection of synchronic gradience and grammaticalization we make some observations on the limits of the intersection (§3.1). In response to the next question about Aarts’s approach to gradience, we examine the extent to which distinctions between different kinds of synchronic gradience can be upheld from both a typological and a historical perspective, with a focus on the latter (§3.2). The final question concerns the relationship between gradience, structural gradualness, and grammaticalization, with a particular focus on reanalysis (§4.1) and analogy (§4.2). §5 is the

3 See also the quotation from Haspelmath (2001: 16539) provided in the first paragraph of the preface to this volume.
4 To some extent this presents the extremes on the formalist – functionalist continuum. Recent research which models on-going change within the minimalist program (e.g. Adger and Smith 2005, Adger 2006) has paid greater attention to the sociolinguistics of change, though even here, the division between “use” and “grammar” remains, such that “syntax is Socio-free and Use-free” (Adger 2007: 700). Similarly, change (and grammaticalization in particular) has been considered within OT by Kiparsky (Forthcoming); for an exploration of historical syntax in stochastic OT, see Clark 2004.
2. Some background

In order to answer the question about the intersection between gradience and grammaticalization, we propose to distinguish gradience as a synchronic phenomenon, and gradualness as a diachronic one. Synchronically, strings can be arranged on continua of categoriality and of grammaticalness. We see diachronic gradualness as a sequence of discrete micro-steps affecting various aspects of the use and structure of a linguistic sign.

2.1 Gradience

One aspect of gradience is that some members of a category are ‘better’ than others. This relates either to Goodness of Exemplar or to Degree of Membership in prototype theory (Denison 2001, 2006, Aarts 2004, 2007a, 2007b, Rosenbach 2006, 2007, Croft 2007). For example, the noun fun has adjective properties in expressions like it was very/so fun (Denison 2001: 127); the partitive sort of (head) has modifier properties in these sort of ideas (Denison 2001: 134). Another aspect of gradience is the fuzziness of boundaries between categories (Denison 2006). Investigating the more complex and heterogeneous category ‘adverb’, and ways in which particular instances may overlap cross-linguistically within the languages of Europe with nouns, verbs, converbs, and adjectives, Ramat and Ricca (1994) point to the difficulty of establishing the relevant sets.

In an attempt to establish criteria for gradience, Aarts (2004, 2007a, b) suggests that there are two different types of gradience, based on morphosyntactic distributional criteria. These will be discussed in more detail below in §3.2. Suffice it now to point out that he distinguishes intracategorial, ‘subsective’ gradience within a category and intercategorial, ‘intersective’ gradience between categories (Aarts 2007a: 97). Subsective gradience can be illustrated by the ‘goodness of fit’ of items within the category of adjectives, for example, their ability to appear in attributive and predicative position, and to cooccur with an intensifier, to be graded, or prefixed by un-. Utter fails on all but the first criterion:

(1) a. an utter disgrace (attributive)
   b. *the problem is utter (predicative)
   c. *very utter (intensification)
   d. *utter/utterer/utterest (gradedness)5
   e. *unutter (un- prefixation) (Aarts 2007a: 106)

Another example is gradience within a single class of verbs (Aarts 2007a: 98-101), from main verb (hope to) to catenative (seem to), semi-auxiliary (have to), modal idioms (had better), marginal modals (dare), and central modals (can) (see Quirk, Greenbaum, Leech, and Svartvik 1985: 137).

While subsective gradience involves only a single category or set of properties, and is assessed in terms of prototypes, intersective gradience involves two categories or sets of properties which may converge “on a cline” (Aarts 2007a: 97). An example of intersective gradience is the distinction between adverbs and adjectives, e.g. some adverbs ‘mimic’ adjectives, as does now in the now generation (p. 136), almost in his almost-victory (p. 138); see also Rosenbach (2007, this volume) on possessives and NN compounds. Likewise, Denison (2006) treats both noun-adjective gradience (very/so fun) and partitive-postdeterminer gradience (sort of) as intersective.

---

5 As Denison points out, utterest occurs in e.g. utterest nonsense/contempt, and should at best be marked with a ‘?’.
2.2 Gradualness

Gradualness refers to the fact that most change involves (a series of) micro-changes, an issue which is sometimes overlooked in considerations of more general patterns of language change. As Brinton and Traugott (2005: 150) observe, although change is sometimes understood (or at least formulated) as $A > B$, studies of gradualness in linguistic change attempt to uncover “the tiny local steps between $A$ and $B$ that the arrow ‘$>$’ encompasses”. Successions of micro-changes may in some cases lead to macro-effects, as Lightfoot (1979) showed. He privileged macro-changes. However, since these do not occur without prior micro-changes, the latter need to be theorized as much as macro-changes. Micro-changes are discrete (Hopper and Traugott 2003) and, as conventionalizations, cognitively abrupt (in a tiny way) for individual speakers\(^6\). However, on the assumption that innovation is not change, only consolidation of an innovation via transfer to a community is, changes at the level of the community are not discrete/abrupt.

Following Timberlake (1977) and Lichtenberk (1991) it has become standard in much of the grammaticalization literature to think of reanalysis followed by actualization,\(^7\) in other words of “the formulation of a novel set of underlying relationships and rules”, followed by “the gradual mapping out of the consequences of the reanalysis” (Timberlake 1977: 141; developed further in e.g. Harris and Campbell 1995, Andersen 2001, 2006). In the functionalist literature a distinction is made between actualization that occurs within the linguistic system and actualization across speakers, spaces, and time. The former, ‘structural gradualness’, is an orderly progression across semantic/pragmatic, lexical and syntactic structures (see Himmelmann’s 2004 three types of expansion: semantic-pragmatic, host-class, and syntactic). It may involve gain or loss of properties one by one (Lichtenberk 1991). Sometimes structural gradualness is regarded as a progression from more marked to less marked status (e.g. Timberlake 1977, Andersen 2001, Fanego 2004). The concept of structural gradualness is not consistent with parametric analysis, and so has not been adopted in most generative frameworks (although appeals to markedness have been, see e.g. Roberts and Roussou 2003). Parametric change is considered ‘abrupt’, and gradualness in this framework is reserved for transition of types and tokens through the community correlated with increased frequency of types and tokens (Kroch 2001, Hale 2007, Roberts 2007: 335). According to Lightfoot, this ‘social’ gradualness is actually outside of the purview of generative historical syntax (e.g. Lightfoot 1991: Chapter 7, 1999: Chapter 4).

‘Reanalysis plus actualization’ has had the unfortunate effect of contributing to a polarization of reanalysis and gradualness, despite Harris and Campbell’s attempt to downplay the contrast (1995: 48). For example, arguing that language change is grammar change and comes about solely through child language acquisition, Lightfoot (1979 and elsewhere) privileges abruptness, understood as saltation, and catastrophic change,\(^8\) all of them associated with reanalysis (and recently with parameter-change) after multiple local micro-changes (e.g. Lightfoot 1991, 1999, Hale 1998, Roberts 2007). But over time the parameters have become smaller (see e.g. Cinque and Kayne 2005), and even though parametric change is regarded as abrupt, the abruptness can hardly be regarded as ‘salutation’.

Another possible reason for the association of reanalysis with saltation is the unintended consequence of the way in which that most ‘clines’ of grammaticalization are represented. Typically they are represented schematically in terms of broad changes (macro-steps) to distinct-seeming

---

\(^6\) We are here assuming that at some point speakers have a new representation of a linguistic expression that they did not have at an earlier time. This may be an abstraction away from neurological processes involved in change, but is necessary if we are to distinguish ‘change’ from ‘emergence’. ‘Emergence’ is ambient and non-discrete, whereas ‘change’ picks out conventionalizations that become part of a community’s repertoire at some approximate point in time.

\(^7\) ‘Actualization’ has to do with “how a language moves from one state to a succeeding state” (Kroch 2001: 726), and is the equivalent of what Weinreich, Labov, and Herzog (1968) identified as the “transition problem”.

\(^8\) While in earlier work Lightfoot referred to “cataclysmic re-structuring” (1979: 78), more recently this has been defined less dramatically as “bumpy discrepancies” between input to a child and output of the child’s mature grammar that have significant systemic effects (Lightfoot 1999: 89).
categories along a cline, such as:

(2) main verb > auxiliary verb > clitic > inflection

\( \text{will ‘to will’ [main verb]} > \text{will [auxiliary]} > \ ‘ll [clitic]} \)

(3) Partitive/Measure Phrase > Quantifier > Degree Modifier

\( \text{a lot of fans is for sale ‘a unit of’ [Partitive]} > \text{a lot of fans are for sale ‘many/much’ [Quantifier]} > \text{a lot happier ‘very much’ [Degree Modifier]} \)

These schemas are, however, “generalizations over changes” (Andersen 2001: 214), and (deliberately) idealize over many things. Above all, since they identify macro-types of change, not individual micro-changes, they obscure the orderly progression of changes across linguistic contexts, such as the host-class expansion (Himmelmann 2004) of auxiliating verbs to new main verb collocates (see e.g. Hilpert 2008). Another factor that these clines obscure is that both the old and new structures coexist in individuals as well as communities, e.g. at the time that be going to was first used as an expression of future it coexisted as a polysemy with the motion-with-a-purpose expression. We may hypothesize that the expressions became homonymous only after a considerable length of time, and after a sufficient number of changes had occurred to allow phonological reduction of the future. When they did so, their relationship to each other presumably became opaque and ultimately completely broken (“divergence”, Hopper 1991).

Micro-steps, then, are ultimately consistent with gradualness, given a theory of continuity over time, and of synchronic polysemy. On the view presented here, gradualness is not an issue of indeterminacy, vagueness, or undecidability (Klamer 2004), although particular instances of potential change may be undecidable, even in context. Nor is it to be understood as an issue of drift (Haspelmath 1998), or monotonic incremental change as Janda implies when he equates gradualness with continuity (2001: 307, questioning how “continuous/gradual” increases in grammaticalization are). It is also not to be equated with slow progression over long periods of time (Bruyn 1996: 39). Changes may occur at different times and at different rates (pace Kroch 1989, 2001). The speed at which some changes come about may in part be a function of other ongoing changes in the language, since any particular change is embedded in the larger context of changes in the language (Weinreich, Labov, and Herzog 1968), whether or not there is already a frequently-used model, or whether contact has occurred. Nor is gradualness “imperceptible” change (Kiparsky 1968: 175).

Because gradualness is discrete and identifiable by small-scale changes in linguistic properties, we can see gradualness as in some ways a diachronic dimension of gradience. However, it is not the dynamicized equivalent of gradience for reasons to be discussed immediately below.

3. How are we to understand the intersection between synchronic gradience and grammaticalization?

Various hypotheses have been put forward about the likelihood that synchronic variation will reflect past history. For example, Bybee, Pagliuca and Perkins (1991) hypothesize that the most inflected expression of futurity is the oldest:

[F]utures with only or predominantly ‘late’ uses are morphologically reduced relative to ‘young’ futures, suggesting a dynamic correlation of generalization of meaning with concomitant reduction in the formal expression of future grams. (Bybee, Pagliuca and Perkins 1991: 47)

Heine and his colleagues have further argued that history can be reconstructed from synchronic

---

9 Hock (1991: 633) notes that the neogrammarians “had recourse to the criterion of ‘imperceptible/gradual’ vs. ‘perceptible/abrupt’”. 
variation (e.g. Heine 2002). However, systemic changes may interfere, and if so the reconstruction will not be supported. Nevertheless, it is a reasonable hypothesis that gradience is the result of small-step changes such as are associated with gradualness. Understanding the limits on and the issues connected with this intersection is crucial.

In what follows we assume that some instances of grammaticalization do intersect with synchronic gradience. The question is which aspects of grammaticalization do so (§3.1), and whether Aarts’s model of gradience can allow such intersection to be accounted for (§3.2).

3.1 How does grammaticalization intersect with gradience?

A basic assumption of grammaticalization, understood from a functional perspective as involving meaning and form, is that at the level of specific changes, the process begins with mismatches between the pragmatics/semantics and form of a construction (in the Construction Grammar sense of a form-meaning pair, Croft 2001), and is manifest in various morphosyntactic ways, including decategorialization (e.g. Hopper and Traugott 2003). Therefore, on first thought, grammaticalization and gradience should intersect easily. But the interface is actually quite problematic, and somewhat minimal, especially where gradience is defined purely or even primarily in morphosyntactic terms.

a) Gradience as defined by Aarts and Denison (for example) is conceptualized in terms of discrete syntactic or of morphosyntactic distributional properties, with little attention to function and meaning, as exemplified by: “Assignment of an element $a$ to a particular class takes place by examining and enumerating the morphosyntactic properties of that element” (Aarts 2004: 21). Furthermore, Aarts privileges syntax as autonomous. However, much work on grammaticalization is conceptualized in terms of degrees of grammatical function. This is on the assumption that there is a continuum between lexical, contentful constructions and grammatical functions (Lehmann 1985, 1995, Hopper and Traugott 2003). Noun, verb, and adjective are on the contentful end, while grammatical ones are on the abstract, procedural end. Grammaticalization shifts a linguistic expression “further toward the functional pole of the lexical-functional continuum” (Haspelmath 1999: 1045), and specifically to function as a marker of “deictic relational structure” (Diewald Forthcoming). This means that only a sub-part of what is considered to be grammaticalization could possibly intersect with gradience as construed by Aarts and Denison.

b) What is often considered criterial for grammaticalization is that there is a sequence of changes. Typically, semantic/pragmatic, morphosyntactic, and morphophonological changes may affect an item that is grammaticalizing. Bybee, Perkins and Pagliuca (1994: 20) refer to this phenomenon as “the dynamic coevolution of form and meaning”, while McMahon (1994: 161) has called it “cross-componential change”, or “a syndrome of related changes”. Denison (2005), Traugott (2007) and Trousdale (2008c), among others, have discussed the development of degree modifiers from ‘NP of NP’ constructions (Denison on sort/kind of, Traugott on a bit/lot, and Trousdale on a hell of a). In each case the decategorialization of the first noun in the ‘NP of NP’ sequence is shown to be a part of a wider grammaticalization process involving a series of pragmatic inferences, semanticizations, syntactic reanalyses and phonological reductions. Most grammaticalization sequences evolve over considerable lengths of time; sometimes the original source is lost (e.g. the original main verb use mot- ‘able to’ that is the predecessor of must (past tense form)), or becomes homonymous with it (e.g. will as main verb and as auxiliary). This means that all the various stages of a particular instance of grammaticalization may not coexist at any one moment in the history of a language, or if they do exist, they may not be considered to be in a gradient relationship to each other, so the link to synchronic variation may not be direct.\(^{11}\)

c) Relatedly, in the grammaticalization literature, the fact that older and newer variants are polysemous (“layered”, Hopper 1991), at least initially, is of key importance. From this perspective,
‘non-central’ members of a gradient cline are assumed to originate in central ones, and can be shown to do so from historical texts. Famously, current ‘core modals’ originated in main verbs that could be inflected, have objects, etc. (Lightfoot 1979, and virtually every study of the history of modals since then). Expressions that are grammaticalizing could be expected to have both core and more marginal uses, i.e. be multifunctional. However, in Aarts (2007a) gradience is not discussed in terms of polysemy, or coexisting patterns of usage, so again there is a mismatch of structures under consideration.

d) A better fit with the concerns of work on gradience can, however, be found in the current shift in research on grammaticalization to membership of categories (e.g. what expressions may become degree modifiers, modals, etc.). Work on polysemy traces the development of the meaning of an expression while keeping form relatively constant, e.g. motion be going to > future be gonna. By contrast, work on what expressions come to serve a particular function traces the development of forms while keep meaning relatively constant. The shift in interest toward patterns arises in part out of focus on patterns of Source and Target changes (Heine and Kuteva 2002), and on constructional change (e.g. Bergs and Diewald 2008), understood as changes in form-meaning pairings. We see either the creation of a new construction (rare) or the gradual reconfiguration of extant constructions (more common), in which membership of a particular constructional ‘set’ (i.e. a series of related constructions at different degrees of schematicity, clustered around a particular node in the constructional network) is involved in expansion and change, through the acquisition of more indexical or procedural meanings (Trousdale 2008b).

e) Again, to the extent that gradience can be seen to be a characteristic of systemic or typological structure (e.g. transitivity is said to be a more flexible construction in English than in German, Taylor 1998), the intersection with grammaticalization will be the stronger. Change does not occur in a systemic vacuum. In particular, specific as well as more general types of grammaticalization may be constrained or expanded because of systemic changes in a language (Lehmann 2004, Fischer 2007). Likewise, it has been suggested that counterexamples to grammaticalization, such as the degrammaticalization of –s genitive in English and Swedish, can be accounted for as residues after radical changes, in this instance, ‘deflexion’ or loss of inflections (Norde 2001).

In sum, at best, only the tip of the iceberg that is grammaticalization intersects easily with synchronic gradience, to the extent that the latter is understood as syntactic and morphosyntactic distribution.

3.2 What insights does grammaticalization offer for assessing the validity of Aarts’s (2007a) claims that there is a significant distinction between subsective and intersective gradience?

In this section we focus on a particular point of Aarts’s argument, namely that there is a distinction between two kinds of gradience, intersective and subsective. We recognize that Aarts’s work is synchronic, and is not intended to reflect or predict change. However, as stated at the beginning of section 2, some kind of connection between synchronic systems and change is logically inescapable. As Aarts (2004: 37) notes, sequences of changes such as are typical of grammaticalization are of particular concern in establishing the relevant properties for assigning membership of a category. There is a danger that a model of gradience which focuses on distribution alone will fail to capture the right generalizations about the nature of grammatical categories if we do not know what the relevant properties for category membership are (i.e. what attributes we associate with particular categories) or whether some of these properties are more important than others (e.g. whether ability to appear in attributive position is a more central characteristic of adjectives in English than the ability to appear in predicative position). This is compounded by the fact that, as Aarts himself recognizes, “languages are not static entities” (Aarts 2004: 37), and is of particular concern when it comes to establishing patterns associated with intersective gradience: intersective gradience is not the outcome of an intersection of categories, but rather of an intersection of properties associated with particular categories.
As noted in §2.1, Aarts understands gradience in terms of discrete morphosyntactic distributional properties. Subsective gradience involves subcategorizations, for example *before, since*, can be regarded as subcategorized according to such patterns as __NP, __S, __0 (i.e. uses as prepositions, conjunctions, and adverbs) (Aarts 2007a: 150-154). Intersective gradience involves overlapping morphosyntactic properties, e.g. both Adj and Adv can serve as modifiers (*then, upstairs*, Aarts 2007a: 136-138). Aarts’s main aim is to come to a position “that allows for gradience, but nevertheless maintains sharp boundaries between categories” (Aarts 2004: 1), in an attempt to find a middle ground between formalists who marginalise gradience to the periphery of grammatical analysis, and functionalists who believe that gradience is central because it is so widespread in grammars; furthermore he suggests that intersective gradience is less common than subsective gradience.

If we wish the two types of gradience to have any reality for change, for example, to reflect or even predict optimal paths of change, then we might expect them to reflect or predict different trajectories. By hypothesis subsective gradience might lead to well-attested change, while intersective gradience might lead to changes that are less widely attested, or that have undergone more complex sets of change. But neither part of this hypothesis is borne out. With respect to subsective gradience, we may ask whether any verb has gone through successive sub-parts of the verbal gradient mentioned in section 2.1: main verb - catenative - semi-auxiliary - modal idiom - marginal modal - central modal. There is no evidence that any verb has done so in English or any other languages; however, the cline cited in (2) above is cross-linguistically well attested. Cross-linguistic diachronic evidence shows that verbs may gradually be assigned more modal-like properties, without necessarily first acquiring properties of a catenative, then properties of a semi-auxiliary, then properties of a modal idiom, and so on. With respect to intersective gradience, we may consider the example of the development of complementizers. One well-known source is demonstratives (Frajzingier 1991, Diessel 1998), and another is verbs of speaking, such as *say* (Deutscher 2000: Chapter 5). They do not appear to have had significantly different histories from the development of auxiliaries that might be attributed to intersective gradience. Furthermore, there is no evidence that *that*-complementizers arose via an adjectival stage; they always appear to have had pronominal properties (as in *She said that: that X*).

Further problems exist within patterns of what would need to be subsective gradience according to the Aarts model, in the development of adpositions and subordinators in Romance and some Bodic languages of Sino-Tibetan (see further Croft 1991, 2007). The grammaticalization of adpositions to subordinators in these languages is not uniform, with formatives in different languages displaying different properties. To this extent, it is questionable whether prepositions and complementizers should be brought together under a single category (an issue which Aarts himself mentions is up for debate). What might be subsective gradience in one language may be intersective gradience in another; and the diachronic data raise further concerns.

Aarts concludes his book by saying that intersective gradience “is arguably far less widespread than is often claimed because many perceived cases are the fall-out of the less than optimal way grammarians have set up their categorial taxonomies” (Aarts 2007a: 242). His attempt to be precise about a structural way to account for gradience has brought many important questions to the fore. However, in so far as the hypothesis of intersective gradience may be tested by evidence from gradualness and from grammaticalization it appears that intersective gradience may itself be the fall-out of a less than optimal way of thinking about gradience. Aarts also claims that it is “paradoxical for adherents of gradience models that their starting point must always be rigid discreteness” (2004: 19). However, the starting point would not have to be rigid discreteness in a constructional approach to grammaticalization. Grammatical constructionalization starts out from a position in which rigid discreteness of formatives is seen as atypical in language, because the primitive is not the formative, but the construction (Croft 2001, 2007). This is made particularly clear when one considers the notions of reanalysis and analogy. We address these issues in the next section.
4. What does the intersection between synchronic gradience and grammaticalization tell us about whether work on grammaticalization needs reanalysis and analogy/extension, or some other mechanism?

In answer to this question, we propose that the gradience that is attested synchronically arises as the result of successive small-step changes resulting from the operation of the well-known mechanisms of reanalysis and analogy (the ‘how’ of change). As Bybee has pointed out:

By postulating a finite set of mechanisms attributable to human neuromotor, perceptual, and cognitive abilities, which interact with linguistic substance in acquisition and in language use, a range of possible language structures and units will emerge. (Bybee 2001: 190)

These structures are hypothesized to have similar properties because “the cognitive and neuromotor capabilities of human are the same across cultures” and “uses to which language is put are similar” (ibid.). The mechanisms operate on change in general, not only grammaticalization, and are motivated by various activities in which speakers engage (the ‘why’ of change, see section 4.3). In historical linguistics mechanisms are usually conceptualized in terms of ‘before’ and ‘after’, i.e. comparison of earlier and later stages, not of the neuronal processes themselves. Hypotheses about the latter are, however, the focus of Fischer (2007).

Before continuing, it will be useful to summarize some of the arguments that have been presented so far regarding the role of reanalysis and analogy in grammaticalization. In 1912 Meillet famously wrote:

Tandis que l’analogie peut renouveler le détail des formes, mais laisse le plus souvent intact le plan d’ensemble du système existant, la “grammaticalisation” de certains mots crée des formes neuves, introduit des categories qui n’avaient pas d’expression linguistique, transforme l’ensemble du système. [‘While analogy can renew details of forms, but usually leaves the structure of the existing system intact, “grammaticalization” of certain words creates new forms, introduces categories that had no linguistic expression beforehand, transforms the system as a whole’]. (Meillet 1958: 133)

Since then the concept of analogy (at that time construed mainly in terms of what is now known as ‘four-part analogy’ and ‘leveling’) has been refined beyond anything that Meillet would have recognized, and the concept of reanalysis as we now understand it has been introduced. Whether reanalysis can be identified with grammaticalization or whether it has anything significant to do with it has been a topic of debate in the last twenty-five years, starting with Heine and Reh (1984), but foreshadowed in Langacker (1977) and Timberlake (1977), and their discussion of reanalysis in morphosyntactic change. While analogy has long been considered an important factor (see e.g. Givón 1991), it has until recently not been at the center of attention in work on grammaticalization. The fundamental importance of extension in most instances of grammaticalization has in the last

---

12 Harris and Campbell (1995: 50) say there are only three basic mechanisms in syntactic change: reanalysis, extension (including analogy), and borrowing. We will not discuss the third of these here. See also Andersen (2008) for a more extensive set of basic mechanisms in morphosyntactic change. Semantic analogues to reanalysis and analogy are conceptual metonymization and metaphorization (Anttila 1989: 141-142, Hopper and Traugott 2003: 84-93).

13 In a move that is different from that of most other researchers on grammaticalization, Haspelmath (1998: 344) claims that “The main mechanism of syntactic change is grammaticalization”. This is because, in his view, grammaticalization is “the gradual drift in all parts of the grammar toward tighter structures, toward less freedom in the use of linguistic expression at all levels”. Such a conclusion not only privileges grammaticalization as unique but appears to assume that it is possible to interpret repeated moves toward dependency as an independent mechanism. If a mechanism is a neurological process, this seems implausible, assuming changes in brain states are discrete; they may occur repeatedly, but each change is a single step.
decade led to significant rethinking of the role of analogy. Several of the issues are discussed from very different perspectives in Lehmann (2004) and in more detail in Fischer (2007), and will not be repeated here.

4.1 Reanalysis

In a definition that is still widely drawn on, at least in the functionalist literature, Langacker identified reanalysis as “change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface structure” (Langacker 1977: 58), i.e. it is covert. He went on to characterize two types of reanalysis. One is “resegmentation”, i.e. boundary loss, boundary creation, and boundary shift. The other is “syntactic/semantic reformulation” (p. 64), which ranges from semantic change to change in agreement patterns. Drawing on Langacker (1977) and also Timberlake (1977), and assuming a version of Government and Binding syntax, Harris and Campbell defined reanalysis as involving change in constituency, hierarchical structure, category labels and grammatical relations in underlying structure (1995: 50) without change in surface structure, including morphological marking, and word order. They regarded rebracketing and restructuring as specific kinds of reanalysis (p. 51). Crucially, they considered reanalysis to depend on “a pattern characterized by surface ambiguity or the possibility of more than one analysis” (ibid., bold original). This in itself may lead to the semblance of synchronic gradience: in the development of quantifiers emerging from ‘NP of NP’ constructions in English (Traugott 2007), it is possible that a hearer was unable to decide whether or not the speaker intended a partitive or non-partitive meaning (e.g. [I bought [[lots] of fans]] vs. [I bought [[lots of] fans]]). In such circumstances, gradience may be seen as a gradual phenomenon, in which speakers and hearers negotiate meaning, matching patterns to existing constructions, or even modifying existing constructions to accommodate particular instances of use. Recently, Lehmann has defined “reanalysis of a construction” more generally as “the assignment of a different grammatical structure to it.” (Lehmann 2004: 162). Under any of these definitions, particularly the latter, provided we allow both semantics and phonology to be included, as Langacker (1977) does, any structural change will involve reanalysis. Since diachronic grammaticalization involves structural change, reanalysis should necessarily be involved. But the question remains in what way it is involved.

The following main positions have been taken regarding the relationship between grammaticalization and reanalysis:

a) Grammaticalization and reanalysis intersect but are independent (Hopper and Traugott 2003, Lehmann 2004). Arguments put forward for their independence include the fact that grammaticalization is unidirectional but reanalysis is not. Lehmann (2004: 164) cites the following factors as reasons for regarding them as independent: “reanalysis does not imply loss of autonomy” or of information, and is not gradual. He also rejects suggestions that grammaticalization is reanalysis on the grounds that reanalysis consists of two stages, whereas grammaticalization is a sequence “‘S1, S2 …Sn’” (Lehmann 2004: 165) (see section 3.1, point b)).

b) Grammaticalization is derivative of (i.e. an epiphenomenon of) reanalysis, which is itself an epiphenomenon of child language acquisition: “the notion of Diachronic Reanalysis is derivative of aspects of the process of language acquisition. Since grammaticalization is derivative of Diachronic Reanalysis, we see that this is a doubly derivative notion” (Roberts 1993: 254). Typical expressions of this position include: “reanalysis (also sometimes extension) is the determining mechanism that explains grammaticalization and without appeal to these mechanisms grammaticalization has no explanatory power of its own … Grammaticalization is always the result of reanalysis” (Campbell 2001: 151, 144). While Roberts (1993: 252) adheres to the concept of parametric change as “a random ‘walk’ through the space defined by the set of possible parameter values”, Roberts and

14 By ‘construction’ Lehmann presumably means syntactic string, not form-meaning pair.
Roussou (2003: 201) suggest grammaticalization can be “reduced … to an instance of parameter change”, upward “along the functional hierarchy” (p. 202). This upward reanalysis accounts for unidirectionality, in their view, and can give rise to new functional material (p. 209); in this sense, grammaticalization involves a subset of types of reanalysis.

c) Reanalysis is largely irrelevant to grammaticalization because it has properties inconsistent with it. For example, Haspelmath (1998: 315) says “pure grammaticalization” should be explained “without reference to reanalysis”. This is because, unlike grammaticalization, reanalysis is abrupt, non-gradual, bidirectional, requires ambiguity in the input structure, and most importantly does not involve loss of autonomy/substance (Haspelmath (1998: 327).16 There is, however, nothing in the definitions provided by Langacker, Harris and Campbell, or Lehmann, to require abruptness. All change is ‘abrupt’ or ‘discrete’ in the sense that it is discontinuous from generation to generation. It is also ‘abrupt’ when a new meaning or structure comes to be represented in a speaker’s brain.

Before concluding this section, we may note that one can only ‘re-analyze’ something that pre-exists, so if a child learns a language and parses a particular string with a new analysis, no ‘reanalysis’ has occurred from the point of view of the learner.17 When the language learner later acquires the ‘older’ form-meaning pairing, then his or her grammar will have been locally reanalyzed in this respect. Like many metalinguistic terms, including ‘language change’, the term ‘reanalysis’ is therefore not accurate in a compositional semantic sense, except in the case of language users who reanalyze their own structures. The non-compositionality of the technical term may have contributed to misunderstandings about its role in change.

4.2 Analogy

Reanalysis has often been construed as operating at least in part syntagmatically (cf. rebracketing), and covertly. By contrast, analogy has traditionally been construed as operating paradigmatically, most obviously in instances of morphological change, where affixal paradigms may be at issue, and overtly (Anttila 1989: Chapter 5). Like reanalysis, analogy is not unidirectional but “Analogue grammatical change … may cooccur and interact with [grammaticalization] in particular historical changes” (Lehmann 2004: 162). By the latter he means that in some individual instances the change seems to have been helped by the prior existence of a model. For example, the grammaticalization of Latin hab- ‘have’ after an infinitive, as in cantare habeo, into a tense inflection, as in Italian canterò, was helped by the prior existence of verbal tense affixes. The presence of models may also account for the “fertility” of certain grammaticalization paths, in some languages, such as Modern Mandarin co-verbs > prepositions (p. 161).

Whereas much work on analogy has focused on individual local changes, and attraction to exemplars, a recurring theme in later work has been the possibility of conceptualizing rule generalization/extension (or, more recently, constraint optimization) as analogy at a higher metalinguistic level of analysis (Kiparsky 1968, Forthcoming). Kiparsky (Forthcoming) suggests a distinction should be made between exemplar-based analogy, which is language-specific and may

15 ‘Pure grammaticalization’ presumably means grammaticalization that does not involve analogy (see Lehmann's contrast between “pure grammaticalization without analogy” (Lehmann 2004: 161) and “analogically-oriented grammaticalization” (Lehmann 2004: 162). In our view, ‘pure grammaticalization’ does not, and indeed cannot, exist if we consider change to be change in patterns of use.

16 The implication that grammaticalization does not involve ambiguity does not appear to be founded in any textual evidence. However, we do not agree with Heine (2002) and Diewald (2002) that pragmatically ambiguous contexts (whether ‘bridging’ or ‘critical’) are a ‘necessary’ prerequisite for grammaticalization, especially when grammaticalization of non-lexical material occurs, e.g. in the grammaticalization of information structure. Nevertheless, where corpus data are available, they show that in many cases of lexical > grammatical change examples in which the new structure is only potentially inferrable clearly do precede unambiguous ones (see Traugott Forthcoming).

17 Andersen (2001: 231, ft. 3) has suggested the term “neo-analysis” or “neanalysis” for novel analyses by language acquirers.
lead to counterexamples to the unidirectionality of grammaticalization, and UG-based constraints-based analogy, which leads to optimization, hence unidirectionality.

If, in the spirit of usage-based grammars, we think in terms of exemplar-based analogy, including extension, grammaticalization involves generalization to greater type- as well as token-frequency, e.g. grammaticalized be going to ‘future’ came to occur over time with more and more verbs, not only action but also state (Bybee 2003; for a statistical account see Hilpert 2008). Himmelmann (2004) sees this kind of structural ‘host-class’ expansion as key to grammaticalization. Here too we may see intersections with synchronic gradience. In the development of hell of a (→ hella in Californian English) as a degree modifier, we see host class expansion from co-occurrence with bare nouns (a hell of a calling) and with premodified nouns (a hell of a short journey) to co-occurrence with bare adjectives (hella good) and adverbs (hella quickly). Co-occurrence with premodified nouns allows for a (semantic and syntactic) gradience in which ‘critical’ contexts occur, allowing for alternative parsings; indeed, only context will determine whether, when a speaker describes something as a hell of a short journey, he means that the journey was unbearable, though short, or whether it was surprisingly short.

Analogy may therefore be understood as a mechanism of change: the earlier structure is restructured to match an existing one, and as a result, the newer one has a new structure. This is reanalysis in the sense of “the assignment of a different grammatical structure” (Lehmann 2004: 162). Fischer (2007) emphasizes that analogical thinking and reasoning (Itkonen 2005) precede much change, a point with which we agree. It is a point made some time ago by Givón, among others:

Almost all creative-elaborative diachronic change in language, be it phonological, morpho-syntactic, semantic or discourse-pragmatic, is in principle analogical. That is, it involves the language user’s recognition – conscious or subliminal – of similarities between two structural or functional contexts. (Givón 1991: 258, emphasis original)

In this sense it is a precondition or ‘motivation’ for much change, but does not constitute change, or even innovation in the individual. Fischer emphasizes that pattern-finding is typical of primates (Fischer 2007: 249), another point with which we agree. But when it comes to an instance of innovation, she assumes that analogy may not involve reanalysis. For example, she says that:

When a construction like I am going to get some water is ‘reanalysed’ from 4(a) to 4(b):

(4) a. [I am going [to get some water]]
b. [I am going to [get some water]]

it in fact joins another token-set. It leaves the construction-type of [V [to INF]] and joins the construction-type of [Aux V]. (Fischer 2007: 145)

This claim is immediately preceded by “reanalysis in fact does not occur in the physical sense of the word” (p. 145). For a speaker who innovates (4b), that structure exists in addition to (4a), and so (4a) has hardly ‘left’ one construction type and joined another: an additional mental representation has been created for the relevant part of the string in (4a). If reanalysis does not occur in a physical sense, neither does leaving one linguistic category for another. In any event, a classic case of reanalysis has occurred: rebracketing without overt surface change. Furthermore, acquisition of membership of a different construction-type is a category change, and hence a case of reanalysis. Rebracketing and re-categorizing have neuronal implications. The speaker now has a different

---

18 (4a) and (4b) appear as (3a) and (3b) in Fischer (2007).
19 By ‘construction’, Fischer means ‘string’, not ‘form-meaning pairing’.
knowledge structure, arrived at by parsing. In so far as a child or other speaker adopts the new structure, they have interpreted strings like *I am going to get some water* in a different way from speakers who do not have (4b). The question “whether there is any meaningful place left for reanalysis as a type of change” (De Smet 2009) seems to be based on the problems mentioned above of the ‘re-’ in the term, the association of reanalysis with abruptness, and neglect of relationship between reanalysis (mechanism) and parsing (motivation).  

4.3 Distinguishing mechanisms from motivations

We assume that it is important, at least for heuristic purposes, to distinguish motivations, preconditions and potentials for change from actual processes of change, even if the distinction is sometimes hard to uphold. We therefore distinguish between parsing and interpretation as a precondition and potential for change, and reanalysis as a mechanism. Parsing is consistent with indexicality and metonymic thinking (see Hopper and Traugott 2003: 92-93, Fischer 2007: 126 on factors involved in reanalysis). The change involved is discrete, but typically the steps are minimal, i.e. structurally gradual. The fact that accumulated micro-changes may in the end lead to macro-changes can be captured by the term ‘macro-reanalysis’.

A parallel distinction can be made with respect to analogy. Analogical thinking is a motivation or precondition for change that needs to be distinguished from the mechanism. We propose to continue calling the former ‘analogy’ and to dub the latter ‘analogization’ in order to index change. Analogical thinking is consistent with notions of iconicity, metaphor, and set-making (Hopper and Traugott 2003: 92-93, Fischer 2007: 126). The mechanism in individual cases is, however, in fact a reanalysis. Hopper and Traugott (2003: 68) already suggested that all analogy may be reanalysis, and Kiparsky (Forthcoming) points out that his theory of UG-based analogy results in the unification of grammaticalization “with ordinary analogy—not just in the trivial sense of classifying them both as instances of reanalysis”. Like all kinds of reanalysis, then, analogization is discrete. This means there is an asymmetric relationship between analogization and reanalysis: all analogization is reanalysis, but not all reanalysis is analogization. Analogy in the sense of analogization is therefore not primary and reanalysis secondary (*contra* Fischer 2007: 329). This is a claim about mechanisms. If we turn to motivations, and prerequisites for grammaticalization, analogical thinking and reasoning (Itkonen 2005) may be more important than has often been recognized, but parsing is also crucial. This in fact would seem to follow from Fischer’s reliance on Anttila’s (2003) “analogical grid”, which has two axes: form and function, contiguity (the indexical axis) and similarity (the iconic axis) (Fischer 2007: 323).

In answer to the question whether grammaticalization needs reanalysis and analogization/extension, therefore, we suggest that grammaticalization is not reducible to any mechanism (see Lehmann 2004). It is the tendency for successive micro-reanalyses to occur, some involving analogization, some not. Sometimes such successive occurrences may take place over many centuries, sometimes they may be fairly rapid. The two mechanisms discussed here, reanalysis and analogization/extension operate independently of grammaticalization. While analogization/extension are far more wide-spread than has often been recognized, reanalysis can be regarded as the dominant mechanism since there is no change without reanalysis.

5. Conclusion

20 De Smet (p.c.) argues, however, that parsing is based on analogical thinking.

21 Fischer (2007) writes of analogy as both motivation and mechanism. She (personal communication) considers that a distinction between motivation and mechanism is not really necessary. But from the perspective of the historian of language, motivation needs to be separated out as it logically precedes change, and involves factors that are largely language-independent.
Although there is a clear relationship between diachronic gradualness and synchronic gradience, the evidence from grammaticalization has shown that the intersection of all three raises many challenging issues. It is true that some instances of synchronic gradience may be the result of grammaticalization – but others are not, e.g. gradience of ‘nouniness’ in cases such as *hold one’s breath, take one’s time* (Ross 1981, cited in Aarts 2007) is the result of lexicalization. It is true that, taken as a whole, grammaticalization is a gradual process – but the micro-steps involved are discrete and therefore abrupt for individual speakers. It is true that some of these micro-steps may give rise to gradient systems at any synchronic ‘slice’ in the development of a particular language – but successive reanalyses may take place either over considerable periods of time, or they may be fairly rapid. We have discussed some of the ways in which gradience, gradualness and grammaticalization construed as functional change intersect. We have defined gradualness as discrete (‘abrupt’) micro-steps in change. We have suggested that gradience is best distinguished from gradualness as a synchronic phenomenon, and we have established some of the ways in which models of synchronic gradience might be refined given the evidence from grammaticalization.

From the perspective of grammaticalization, it would appear that distributional gradience, without a distinction between subsective and intersective gradience, can be matched with structural aspects of grammaticalization at particular instances in the history of an expression. But grammaticalization involves far more, and to account for semantic as well as morphosyntactic factors a variationist and constructional approach to gradience may be preferable (see Croft 2007). The importance of (synchronic and diachronic) variation is critical here. Variation over time involves the emergence of grammatical constructions: a gradual, global process, but one which involves a series of discrete local micro-reanalyses. Two major points fall out from this. First, while we agree that analogical thinking is an important motivation for change, we see the mechanism of analogization as a process involving discrete reanalyses to other constructional types. Second, because the emergence of constructional types is a gradual phenomenon involving changes at all levels of the construction, accounts of synchronic gradience will be enriched by a consideration of the “meaning or function that the constructions and formative encode” (Croft 2007: 422). This is not to suggest that morphosyntactic distribution is irrelevant, merely that it is only part of the appropriate methodology for establishing typologically consistent patterns of linguistic structure.

Ultimately, as work on the gradience-grammaticalization interface is carried out, gradience could be an ideal testing ground for hypotheses about the expressions that are most likely to undergo grammaticalization, and perhaps more critically, the way in which grammatical constructions emerge. Both grammaticalization and gradience could be sites for gaining a better understanding of the “step-wise acquisition of properties” (Denison 2006: 300) that is so fundamental to micro-changes, while at the same time the gradual sequence of changes involved in grammaticalization could help us to formulate some of the principles involved in macro-changes.

References


