

How gradual change progresses

The interaction between retrievability and innovation

[AUTHOR]

[AFFILIATION]

Abstract: Grammatical changes are typically gradual. That is, a change affects different grammatical contexts at a different time or pace. This gradualness can be explained by assuming that when one grammatical context is affected by a change, this facilitates extension of the change to other related contexts. In gradual changes, then, each step triggers the next. This paper provides evidence to support such a ‘cascade model’ of gradual change by examining its underlying mechanism. It is hypothesized that as an expression is conventionalized in one grammatical context, its retrievability improves, so that it also becomes more easily available in different yet closely related (analogous) grammatical contexts. The change under study is the development of the noun *key* into an adjective. The change is gradual, starting with increased productivity of compounds with *key* as specifying element (as in ...*the key question of licensing*), leading later to ‘debonded’ uses (as in *this figure is a key one in the argument*), and clearly adjectival uses (as in ...*discretionary awards, which are so key to what we are considering*). Using data from Parliamentary debates in the British Houses of Parliament, different predictions are tested that are generated by the cascade model. If innovative constructions are more likely to emerge under conditions that improve the retrievability of analogically related constructions, this should be reflected in effects associated with entrenchment and priming. Entrenchment effects are tested against variation between individual speakers and across speakers' lifetimes. To test priming effects, the grammatical behaviour of *key* is compared across contexts that do or do not prime *key*, either through earlier uses of *key* in discourse, or through collocates of *key*. The predictions drawn from the cascade model are largely borne out. The evidence supports causal connections between the steps of long-term gradual change, as assumed by the cascade model.

Keywords: adjective; cascade model; debonding; entrenchment; gradualness; grammatical change; individual variation; *key*; noun; partial sanction

1 Introduction

One way or another, grammatical change is nearly always a step-wise process. Innovative forms spread gradually through the grammar of a language, occurring earlier or establishing themselves faster in some grammatical contexts than in others (Aitchison 1991; Harris & Campbell 1995). A few examples suffice to make this clear. Lee (2011) finds that English *much* is developing from a predominantly positive polarity item into a strictly negative polarity item. She reports that the change takes place faster for adverbial uses of *much*, as in (1a), than for pronominal uses, as in (1b).

- (1) a. Mr. Bond did not worry *much* about Pat (1856, COHA)
b. “They didn’t bring *much*,” he said. (1875, COHA)

Chappell (2008) discusses the development of the verb *shuō* ‘say’ into a complementizer in Beijing Mandarin. She suggests that the complementizer use first combined with perception and speech act verbs and later also with cognition verbs, such as *juéde* ‘think, feel’ in (2) (cf. Güldemann 2002).

- (2) wǒ zǒngshì juéde shuō, shēnghuó-lǐ quē-le diǎnr shénme
I always feel say life-in miss-PFV little something
‘I always feel that there is something a little lacking in my life.’ (quoted from Chappell 2008: 84)

[AUTHOR & OTHER] (2013) describe the development of Dutch *wie weet* ‘who knows’ from a clausal matrix, as in (3a), into an epistemic adverb, meaning ‘maybe’. As an adverbial, *wie weet* first appeared in the clausal periphery, as in (3b), then in clause-initial position, as in (3c), and finally in other clause-internal positions, as in (3d).

- (3) a. [...] en *wie weet* of hunne dikwijls herhaalde argumenten niet eenigen indruk op 's Konings gemoed gemaakt hebben. (1841, CHK)
‘... and who knows if their oft-repeated arguments have not left some impression on the King’s mind?’
b. Maar *wie weet*! De tyd vermag alles, en dien vooruit te loopen, zou alles bederven. (1874, CHK)
‘But who knows! Time can do anything, and to run ahead of it would spoil everything.’
c. Enfin, *wie weet* gebeurt het nog eens. (1907, CHK)
‘Anyway, maybe it comes to pass sometime.’
d. Ik knik, opgelucht dat hij iets van zichzelf laat zien, iets dat ik herken en op grond waarvan er *wie weet* nog iets als een band kan ontstaan. (2002, TNC)

‘I nod, relieved that he reveals something of himself, something I recognize and on the basis of which something like a connection could perhaps still develop.’

Numerous other examples could be cited from the literature.

So far, the gradualness of grammatical change has predominantly been interpreted in terms of a *reanalysis-and-actualization model* (Langacker 1977; Timberlake 1977; Harris & Campbell 1995). That model holds that an expression is assigned a new underlying structure in ambiguous contexts (reanalysis), in response to which its grammatical surface behaviour gradually changes (actualization). For example, on this view, Dutch *wie weet*, illustrated in (3) above, was first reanalyzed as an epistemic adverb in contexts that allowed the new analysis. Consequently, it picked up the full behaviour expected of epistemic adverbs, spreading to new contexts incompatible with the old analysis. The model is problematic, however (Haspelmath 1998; Fischer 2007; [AUTHOR] 2013a). Even though actualization is recognized as gradual, there is nothing in the reanalysis-and-actualization model to explain the gradualness. For example, if it is an epistemic adverb, then why does adverbial *wie weet* still resist clause-internal uses other than in clause-initial position? In fact, since the reanalysis-and-actualization model interprets reanalysis as abrupt, the gradualness of actualization is at odds with it.

An alternative model holds that the steps of change, rather than being independently caused by a prior reanalysis, are caused by one another (Tabor 1994; [AUTHOR] 2012). I will here refer to this model of grammatical change as the *cascade model*. According to the cascade model, the likelihood of any given (step of) change is determined by the resemblance of its outcome to already established constructions (see also Naro 1981; Denison 1986; Petré 2012; Rosemeyer 2014). This gives potential new steps of change different likelihoods, which explains why they do not occur all at the same time. The reason change can progress beyond the first most likely steps is that each step taken produces new analogical models to base subsequent changes on, thereby shifting the likelihoods for potential new steps. The history of Dutch *wie weet*, as illustrated in (3) above, can be reinterpreted in terms of the cascade model. As a matrix clause, *wie weet* developed epistemic meaning. Its subsequent formal development was determined by the similarity relations holding between *wie weet* and the epistemic adverbs to which it was analogically attracted. *Wie weet* extended to peripheral contexts first, where it could still pass for an elliptical matrix clause. It then spread to clause-initial position, where it closely resembled both its well-established use as a clausal matrix and its newly established clause-peripheral use. As the clause-initial use emerged, other clause-internal uses received additional analogical support and eventually appeared too. As such, each step of the long-term development is a response to previous steps through local analogies. The cascade model does not have the teleological flaw inherent in the reanalysis-and-actualization model, it is compatible with gradualness and it can predict the specific sequence of steps in a given gradual change ([AUTHOR] 2012). Moreover, the cascade model can be extended to changes that do not involve syntactic category change, such as lexical diffusion ([AUTHOR] 2013b; Rosemeyer 2014).

Underlying the cascade model there can be assumed a simple dynamic at work: shifts in what is acceptable change also what is conceivable. By default, innovations are intrinsically unlikely. That is both because speakers avoid deviation from convention, and because for the linguistic coding of experience they will automatically home in on conventional means first. What is conventional will be better entrenched and constitute a more or less automated mental routine (Langacker 1987: 57-59). In mental retrieval such routines will typically outcompete any less conventional means of expression. Grammatical innovations can only happen when a speaker defeats the odds by retrieving and selecting an expression despite its being unconventional in a given grammatical context – what is called 'partial sanction' in Langacker (1987: 71). There are at least two factors that can shift the odds in favour of the unconventional. First, an unconventional expression will be likely if its deviation from convention is so minimal as to be (almost) undetectable – hence the role of analogy in gradual change. Second, this effect will be the stronger, the more accessible the analogically related conventional expressions are in mental retrieval. This will eventually depend on how well-established they are. To stick with the example of adverbial *wie weet*, the more conventionalized the clause-initial use becomes, the stronger the analogical support it offers to other clause-internal uses. Arguably, then, new and unconventional coding solutions can arise to the extent that speakers have easy mental access to any analogically related conventional coding solutions. In processing terms, one could say that the retrievability of associated conventional patterns 'saves' computational effort, which can be 'spent' again in the creative assembly of a new unconventional pattern. This interaction between the conventional and unconventional should repeat itself with each step of a gradual change.

It is this hypothesized dynamic that the present paper focuses on. The dynamic predicts that the more readily accessible a conventional use of an expression is, the better its chances of being used also in similar but unconventional ways. If the hypothesis holds, new uses of an expression should show up under circumstances that facilitate the expression's retrieval. As already hinted at, conventionalization links naturally to Langacker's (1987) notion of entrenchment, which is associated with discourse frequency. At the same time, retrievability can also be temporarily boosted by priming, which is primarily associated with recency. So the general prediction can be split out into two more specific ones. First, an innovation should become more likely as one of the patterns from which it receives analogical support becomes more firmly entrenched. This prediction will here be tested against individual variation, with frequency differences in the usage of individual speakers taken to reflect different degrees of entrenchment (cf. [AUTHOR] *forthc.*; Schmid & Mantlik *forthc.*). Second, an innovation should also be more likely when one of its analogical support patterns has recently been activated by priming. This prediction will here be tested against the effect of direct priming in discourse and indirect priming through collocational associations.

The change that is to be subjected to these predictions is a relatively simple one, the development of English *key* from a noun into an adjective (Denison 2001). In what follows, the development is examined in detail, drawing on a data set collected from the Hansard Corpus. Section 2 below pre-

sents some necessary background on the development of *key* into an adjective. Section 3 discusses the corpus data. Section 4 describes the different procedures for testing the hypothesis and their results. The significance of the findings is discussed in Section 5. The analysis of the data largely supports the predictions formulated above. Innovative uses of *key* are found to typically occur under conditions that improve the retrievability of already-established uses of *key*. Innovations are more likely to occur in individual speakers who frequently rely on conventional uses of *key*. Innovations are also more likely to occur when conventional uses of *key* have been primed. This is consistent with the mechanism hypothesized to underlie gradual grammatical change, adding further support to the cascade model of change.

2 Background

The development of English *key* from a noun into an adjective has already received ample attention in the literature (Denison 2001, *forthc.*; [AUTHOR] 2012; Vartiainen 2013; [OTHER & AUTHOR] 2014). The beginning of the change has been situated in the second half of the twentieth century (Denison 2001; [AUTHOR] 2012; Vartiainen 2013: 175-176). The change is illustrated by the examples in (4). *Key* in (4a) functions as first element of what is in all likelihood a noun-noun compound, *key phrases*. In other words, *key* is probably a noun here. In (4b), *key* is separated from *observations* by an intervening adjective, *experimental*. Since the English noun phrase normally has adjectival premodifiers placed before nominal premodifiers (Denison 2001), the ordering in (4b) suggests that *key* is an adjective. It certainly is an adjective in (4c), where *key* is used with *very*, an intensifier exclusively found with adjectives. *Key* must also be an adjective in (4d), where it is used predicatively and without the determiner a count noun would normally require in this position.

- (4) a. The proposed wording of the possible agreement was given to Dr. Adenauer with certain *key phrases* in blank. (1952, quoted from [AUTHOR] 2012: 623)
- b. Therefore, we shall start our description of the behavior of electric charges in motion by summarizing the *key experimental observations*. (1961, *Ibid.*)
- c. He alienated a lot of very *key political players* in this town. (1991, *Ibid.*)
- d. Her confirmation was *key* because symptoms like the kind I had can be caused by other factors, too (1991, *Ibid.*)

The existing literature leaves little doubt that the development of *key* is a gradual process (see in particular Denison 2001 and [AUTHOR] 2012). Although finer-grained distinctions are possible, for present purposes, three major stages will be distinguished. First, *key* became increasingly productive as first element in noun-noun compounds. This stage will be referred to as the *compounding stage*.

The compounding stage is illustrated in (5) (or 4a) above. Note that *key*-compounds may consist of more than two nouns, as in (5b).

- (5) a. [...] the Government's desperate desire to shield themselves from effective scrutiny in the first *key* hours [...]. (1996, HC)
- b. [...] not a *key* election pledge, but an important pledge. (1999, HC)

Later, *key* began to appear in constructions that are not exclusively adjectival but much more strongly associated with adjectives than with nouns. While these constructions indicate that the syntactic tie between *key* and its head is loosening, they still marginally allow an analysis of *key* as a noun. This stage will be referred to as the *debonding stage* (after the concept of ‘debonding’ introduced by Norde 2009 and applied to noun-to-adjective shifts in [OTHER & AUTHOR] 2014). Constructions characteristic of the debonding stage include *one*-substitution, as in (6a), and the use of *key* before a premodifying adjective, as in (6b) (or 4b above).

- (6) a. The year 1991 was the *key* one, when the modern scheme came into operation. (1996, HC)
- b. [...] access to most of the *key* official documents in the Sandline affair. (1999, HC)

Finally, *key* adopted behaviour that is exclusively adjectival. This is the *adjective stage*. Constructions characteristic of the adjective stage include adverbial premodification, as in (7a) (or 4c above), coordination to other adjectives, as in (7b), predicative use, as in (7c) (or 4d above), or any combinations of these.

- (7) a. NATO [...] is playing an equally *key* part in building the peace. (1998, HC)
- b. It is true to say that we are setting some *key* and important issues on one side. (1995, HC)
- c. When he elaborated on those special circumstances, one seemed to be *key*. (1997, HC)

As is typical of gradual change, the steps of change overlap. This holds also for the stages in the development of *key*. The increase in the productivity of *key*-compounding continues during the debonding and adjective stages. Likewise, debonded uses continue to grow in frequency during the adjective stage. By way of summary, Figure 1 visualizes the different uses of *key*, as defined above, highlighting the analogical relations holding between them and placing them in the larger constructional networks of noun phrases and adjective phrases under which they resort.

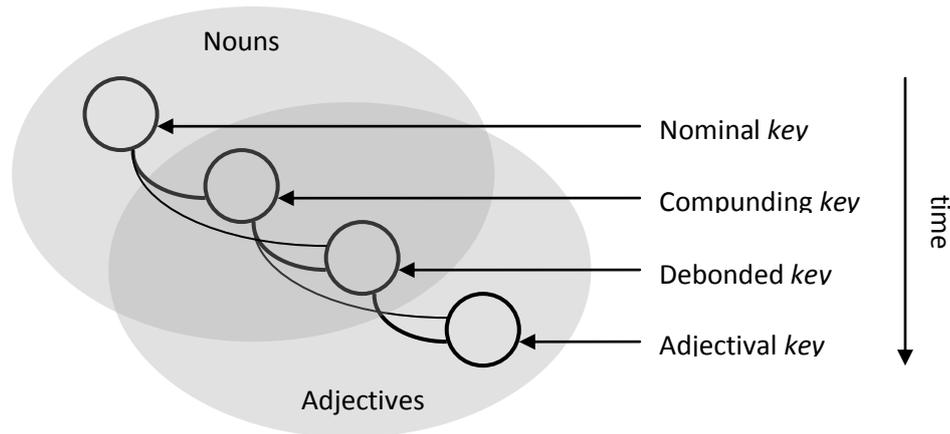


Figure 1. Analogical relations between the developmental stages of *key* and their place within the broader categories of nouns and adjectives.

3 Data

Data on the history of *key* were drawn from the Hansard Corpus (HC). The corpus, compiled by Marc Alexander, is made up of material available from the online Hansard Archive.¹ It contains transcripts of the Parliamentary debates that took place in the British Houses of Parliament, covering almost the entire nineteenth and twentieth centuries. Data on *key* were collected from the transcripts dating from the 1970s, 1980s and 1990s, which is the period that roughly coincides with the onset of the debonding and adjective stages in the history of *key* (see Section 2 above and 4.1.1 below). More details on data extraction procedures are given in the following sections, as they differ from case study to case study.

There were at least two good reasons for using the Hansard Corpus here. One major advantage of the Hansard Corpus is that it allows tracking of individual speakers over time, with very extensive amounts of data available for the more active Members of Parliament. This property of the corpus was exploited in the case studies reported in Section 4.1 below. The other important advantage of the data is that it approximates the structure of conversational interaction. More precisely, Parliamentary debates are organized into conversational turns. Members of Parliament, when taking the floor, typically respond to a previous turn, which we can assume they must have listened to. It is this property of the corpus that was taken advantage of in the case study on priming effects reported in Section 4.2.1 below.

The spoken data represented by the Hansard Corpus differ fundamentally from the speech language users produce in spontaneous conversation. Probably, many Parliamentary interventions have been prepared in writing, perhaps sometimes by others than the speaker they are attributed to. It is also

¹ See <http://www.hansard-archive.parliament.uk/>.

obvious that the transcripts of Parliamentary debates have been heavily post-edited. They lack the hesitations, false starts, inconsistencies and repetitions that are characteristic of spontaneous speech and that are inevitable even in a well-prepared delivery. However, in light of the opportunities the data offer, these features are only minor drawbacks. While the ‘layered’ character of the data may have obscured the tendencies described in the following discussion, it could not have magnified them. After all, it is unlikely that the differences between individual speakers, or the priming effects found across turns result from any ghost-writer activity or editorial intervention.

4 Findings

The following sections apply the predictions made by the cascade model to the history of *key*, testing them on data drawn from the Hansard Corpus. Section 4.1 deals with the effects of entrenchment. Section 4.2 discusses priming effects.

4.1 Entrenchment

The cascade model predicts that innovative patterns should benefit from the entrenchment of any similar-looking patterns. As *key*-compounds become better entrenched, the extension of *key* to debonded and adjectival uses is expected to become easier. Likewise, increased entrenchment of debonded uses should facilitate the emergence of adjectival uses. Taking discourse frequency as a proxy to entrenchment, these predictions are borne out by the long-term development of *key*, with its specific sequence of stages (see Section 2 above and [AUTHOR] 2012). However, there is something of a logical leap in explaining a social phenomenon (the development of *key* in English) by means of a psychological one (entrenchment and analogy at work in individual speakers) (cf. Schmid & Mantlik forthc.). Or, as Bergs puts it:

any claim about cognitive, universal or typological determinants of linguistic change need not only hold for the level of the speech community or its subgroups, but also for a substantial number of speakers in isolation, if it wants to reflect reality. (Bergs 2005: 5)

Therefore, more convincing evidence in support of the hypothesis would be obtained if the prediction also holds for individual variation. For example, speakers differ from each other in the extent to which they rely on *key*-compounding. That is straightforwardly reflected in their usage frequencies. Assuming again that those usage frequencies can serve as a proxy to entrenchment, they should predict which speakers are most likely to produce the more innovative debonding and adjectival uses. On this logic, synchronic variation across individuals is to respect the pattern of diachronic change observed in the overall population. More precisely, speakers’ behaviour should follow a probabilistic implicational

hierarchy, such that behaviour characteristic of a diachronically later stage is likely to be found only in speakers displaying the behaviour associated with diachronically earlier stages (cf. Paunonen 1976). Indeed, a failure to confirm this could falsify the cascade model. If the steps of a gradual change truly follow from one another, it is expected that an individual cannot adopt a (diachronically) later step if they have not adopted the earlier step, even when the later step is already available in the larger population.

In order to study the behaviour of individual speakers, those speakers were selected from the Hansard Corpus who produced at least 300,000 words of running text over one decade (i.e. the 1970s, 1980s or 1990s). Speakers were selected only from the House of Commons transcripts, where speakers are consistently identified by their full name. A total of 153 different speakers matched the criterion. Of these, 32 produced over 300,000 words of text in two consecutive decades and six did so in three consecutive decades. This gives us the split up of speakers over decades reproduced in Table 1, with accompanying total word counts per decade.

Table 1. Division of data over speakers and decades.

Decade	Speakers	Word count
1970s	85	36,239,761
1980s	75	33,160,674
1990s	53	22,016,482

From the subcorpus described in Table 1 all instances of *key* were extracted and linked to the individual speakers who produced them. The examples were manually annotated, sifting out noise and distinguishing between nominal, compounding, debonded and adjectival uses. Figure 2 gives the normalized frequencies of the different uses of *key* centrally at issue here. In the aggregate data for the period, debonded and adjectival uses never rise above a frequency of 0.6 per 100,000 words. Compounding uses are much more frequent, on average about 18 times more frequent than debonded uses, which in turn are about four times more frequent than adjectival uses. This is in line with the diachronic sequence of stages discussed in Section 2 above. The frequencies justify treating the debonded and adjectival uses of *key* as (relatively) unconventional and innovative in the period under study, while the compounding uses are treated as (relatively) conventional and established.² Finally, it should be clear that even the large corpus used here contains rather few of the innovative debonded and adjectival uses. To compensate for that, the case studies in this and the following sections either collapse the distinction between debonded and adjectival uses, or else collapse that between the three different subperiods.

² Denison (2001), writing after the period under study here, reports that (at least) adjectival uses are rejected as ungrammatical by some speakers. This further justifies their treatment as innovative in the 1970s to 1990s.

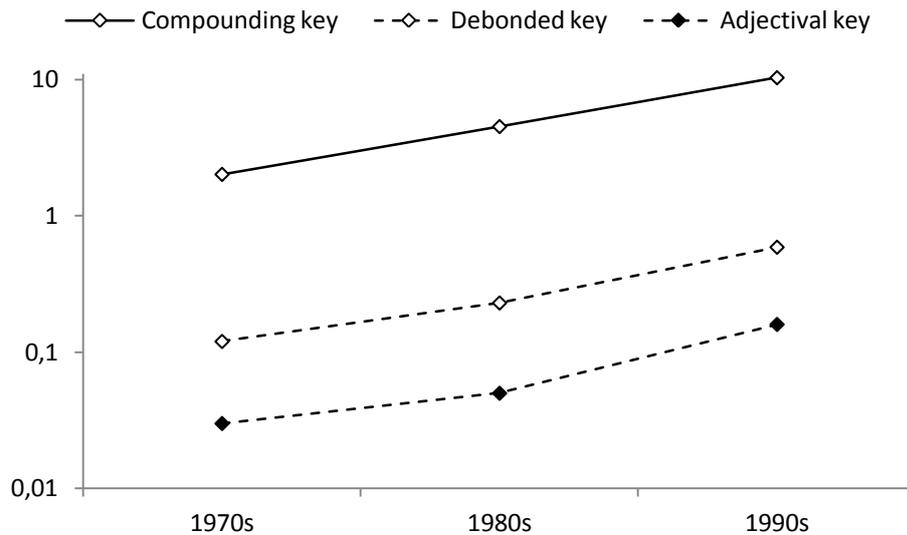


Figure 2. Normalized frequency of *key* in compounding, debonded and adjectival use (frequencies per 100,000 words; logarithmic scale).

Using the annotated data set, a usage profile was determined for every speaker in the corpus. On this basis, it could be investigated whether individuals' synchronic behaviour reflects the diachronic trends in the population. Figure 3 plots individual speakers with respect to their advancement on the different stages of the historical development of *key*.³ Specifically, their use of *key*-compounding (X-axis) is set out against their use of debonded and adjectival *key* (Y-axis). It is found that individual speakers' usage profiles reflect the diachronic order of events. The likelihood for a speaker to use debonded or adjectival *key* increases with their reliance on compounding *key*. The trend is seen in all three decades ($R^2 = 0.26$ for the 1970s; $R^2 = 0.52$ for the 1980s; $R^2 = 0.27$ for the 1990s). This confirms the idea that stronger entrenchment of *key*-compounding will make it easier for speakers to select *key* in the similar, but comparatively unconventional debonded and adjectival contexts. While correlation cannot prove causality, it is at least in line with the interdependence assumed to exist between the steps of a long-term development.

³ The X-axis of Figure 2 uses a logarithmic scale. Speakers who did not produce any *key*-compounds could therefore not be plotted. There are 12 such speakers. None of them produced an instance of debonded or adjectival *key*.

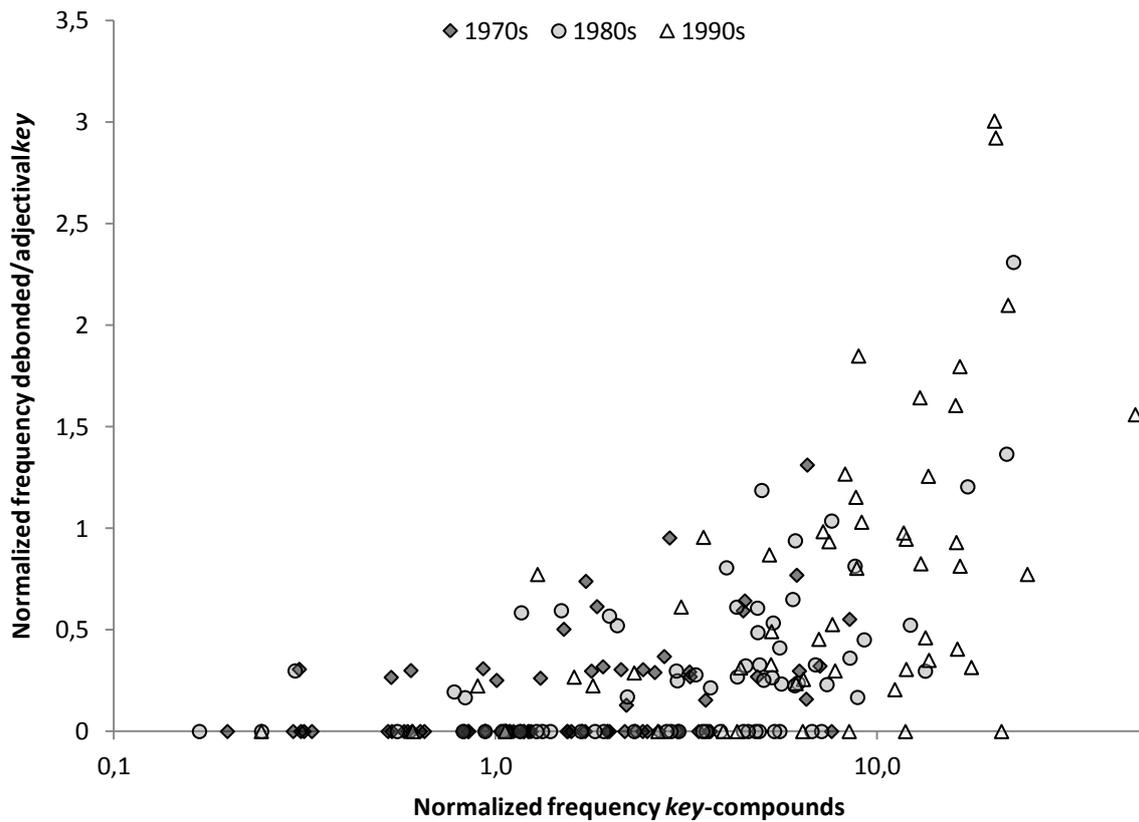


Figure 3. Speakers' use of debonded/adjectival *key* as a function of their use of *key*-compounding.

To corroborate this interpretation, a number of additional aspects of individual variation were investigated. First, there should also be a correlation between speakers' use of debonded *key* and their use of adjectival *key*. Because the incidence of those two constructions is much lower in the data set, the different decades were collapsed. Next, simple distinctions were made (i) between speakers who do or do not display debonded *key* and (ii) between speakers who do and do not display adjectival *key*. This produces four speaker types, defined by the variables [\pm debonded *key*] and [\pm adjectival *key*]. The distribution of the four types in the data set is visualized in the mosaic plot in Figure 4. The plot shows that if speakers use debonded *key* they are also more likely to use adjectival *key* (Pearson's χ^2 9.49, $p = 0.002$; $\phi = 0,211$). The least common speaker type is the one that shows evidence of adjectival *key* but not of debonded *key* – i.e. the speaker type whose behaviour does not reflect the diachronic trend in the population. This is consistent with the cascade model, which assumes that availability of the debonded use facilitates extension to the adjectival use.

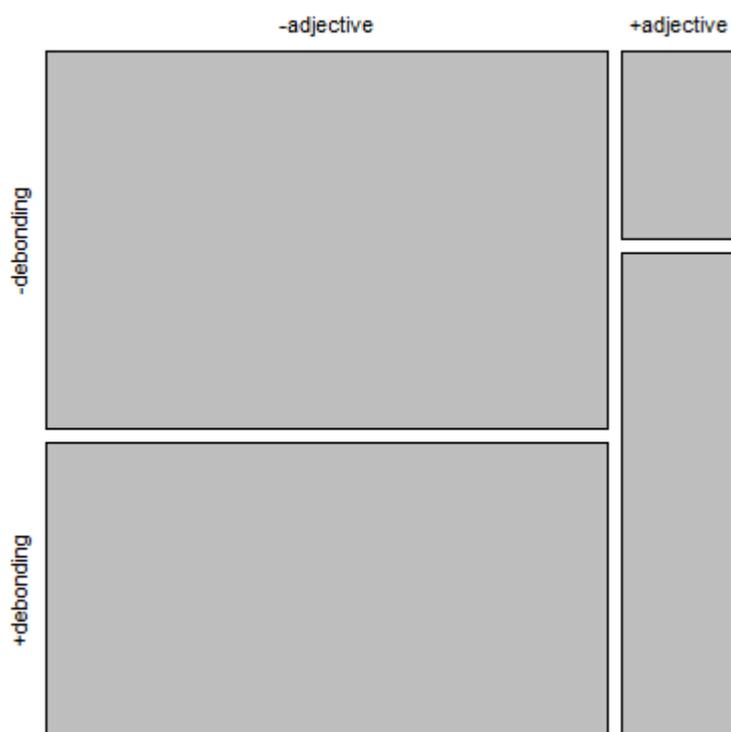


Figure 4. Distribution of speakers (i) with and without debonded *key* and (ii) with and without adjectival *key*.

Second, if the cascade model holds, new steps of a long-term development depend primarily on the immediately preceding step and less so on any earlier stages. In the cascade model, this is what accounts for the observed order of stages in a development. With respect to *key*, it can be expected that for debonded and adjectival uses to arise, the compounding uses were a necessary transitional step. In contrast, the free noun *key* would offer less or no analogical support to debonded and adjectival uses. Individual variation is consistent with this. Instances of (non-literal) free noun uses of *key*, as in (8), were counted per individual speaker and compared to the frequencies of *key*-compounds and debonded or adjectival *key*.

(8) I start with a subject that I have often mentioned before because it is the *key*. (1977, HC)

It was found that, across individuals, the normalized frequency of the free noun *key* positively correlates with the normalized frequency of *key*-compounds ($R^2 = 0.19$ for the 1970s; $R^2 = 0.26$ for the 1980s; $R^2 = 0.25$ for the 1990s) but not or less so with the normalized frequency of debonded and adjectival *key* ($R^2 = 0.04$ for the 1970s; $R^2 = 0.07$ for the 1980s; $R^2 = 0.13$ for the 1990s). This supports the status of *key*-compounds as a transitional stage between the noun and the adjectival *key*. Speakers who frequently use the free noun *key* are more likely to produce *key*-compounds. Speakers who fre-

quently use *key*-compounds are more likely to produce debonded and adjectival *key*. But there is no direct connection between the noun *key* and debonded and adjectival *key*.

Third, the assumed interdependence of the steps of change should also be visible over speakers' lifetimes (cf. Petré & Van de Velde 2014). As pointed out above, for 38 speakers, the data set contains observations over more than one decade. These speakers were divided into four types, depending on whether or not their use of *key*-compounding increased over their lifetimes and whether or not their use of debonded and adjectival *key* increased over their lifetimes.⁴ The mosaic plot in Figure 5 shows how the four speaker types distribute. There is a strong correlation between increasing use of *key*-compounding and increasing use of debonded and adjectival *key* ($p = 0.007$, using a Fischer's Exact Test; $\phi = 0.45$). That is, speakers whose use of *key*-compounds increases over time are also likely to show an increase in their use of debonded and adjectival *key*. Conversely, speakers whose use of *key*-compounds does not increase are unlikely to show an increase in their use of debonded and adjectival *key*. The speaker type whose behaviour goes against the predictions of the cascade model (increasing use of debonded/adjectival *key*, despite decreasing use of *key*-compounding) is also the rarest.

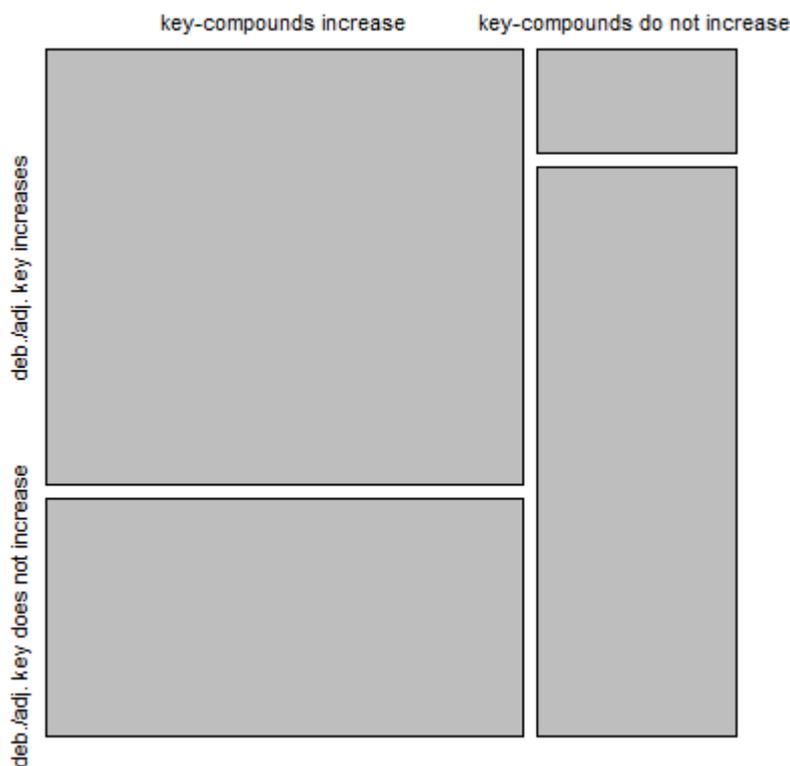


Figure 5. Distribution of speakers (i) whose use of *key*-compounds does or does not increase and (ii) whose use of debonded/adjectival *key* does or does not increase.

⁴ The six speakers represented over all three decades were counted twice, once for the transition from the 1970s to 1980s, and once for the transition from the 1980 to the 1990s.

In sum, variation both between and within individual speakers is consistent with the predictions of the cascade model. Speakers tend to adopt behaviour belonging to the more advanced stages of change only to the extent that they display behaviour belonging to the (immediately) preceding stages. The interdependencies between steps of change revealed here do not prove causal relations, but they are consistent with the assumption of causality.

4.2 Priming

An expression that is strongly entrenched enjoys easy retrievability on a more or less permanent basis. Priming, by contrast, causes a temporary surge in retrievability. The effect has been extremely well documented under experimental conditions. But in natural language use, too, there is good evidence that priming influences speakers' textual choices, for instance in the form of textual persistence (Weiner & Labov 1983; Szmrecsanyi 2006). In light of the research hypothesis, it is expected that priming also facilitates an expression's selection in unconventional contexts.

In the development of *key*, such priming effects may manifest themselves in different ways. Trivially, *key* is likely to prime itself. A speaker should more readily select *key* in their own discourse turn if a previous speaker did so in theirs. The more interesting question, however, is whether conventional uses of *key* prime innovative uses. If so, this would again support the link between increased retrievability and innovation. In what follows, tentative evidence is presented that prior occurrences of conventional *key* facilitate innovative uses of *key* further down in discourse (Section 4.2.1). Next, another type of priming effect will be considered. It has been argued that expressions are primed by their collocates (Hoey 2005). The relation could also be couched in terms of 'cotextual entrenchment' (Schmid & Mantlik *forthc.*). Although the link between collocational and priming relations still awaits more empirical evidence (Durrant & Doherty 2010), it here provides another testable prediction. Specifically, it is expected that *key* is more likely to be used innovatively if it has been primed by other contextual elements it is collocationally associated with. Evidence to that effect is presented below (Section 4.2.2).

4.2.1 Priming across discourse turns

To test whether priming facilitates innovation, the corpus data were manipulated to distinguish between a primed and non-primed condition. For the primed condition, the Hansard Corpus (this time including also the material from the House of Lords) was searched for all instances in which *key* is used in two immediately consecutive turns.⁵ For the 1970s, only 137 such contexts were found, which

⁵ Priming could also have been investigated within the same speakers' turn, but this could have confounded the priming effect with the entrenchment effect discussed in Section 4.1 above. The disadvantage of the across-turns

is not enough to pursue the analysis. The data sets for the 1980s (n=412) and 1990s (n=1,082), however, were manually analyzed to identify cases in which the first *key* was a *key*-compound. Within this subset, the second *key* in the adjacency pair was then analyzed to check the incidence of debonded or adjectival *key*. For the non-primed condition, control sets were collected for the 1980s and 1990s, each containing 1,000 randomly selected instances of *key* not preceded by another instance of *key* within at least 20 turns. These, too, were manually analyzed. Recall here that the point is not to prove that there is priming in the data: in the primed condition, priming is assumed.⁶ Rather, the question is whether *key*, when primed, behaves more innovatively.

Evidence that this is so is found for the 1980s, as shown in Figure 6. When preceded by a *key*-compound in the previous turn, *key* is about twice as likely to be used in debonded or adjectival constructions compared to the control set. The effect is very small but it is significant (χ^2 4.81, $p = 0.028$; $\phi = 0.08$). The finding again supports the idea that innovations thrive under conditions improving the retrievability of their analogical model.

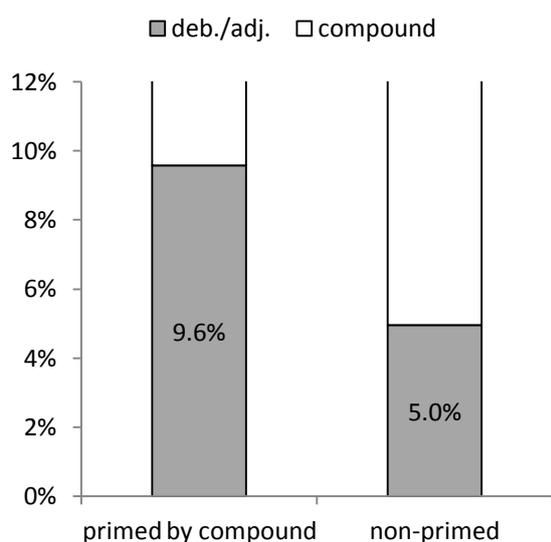


Figure 6. Incidence of debonded/adjectival *key* vs. *key*-compounds following a *key*-compound in the previous turn (1980s).

While this is in line with the hypothesis, the finding needs to be qualified in two ways. First, the finding is not confirmed for the 1990s (χ^2 0.13, $p = 0.71$; $\phi = 0.01$). It is possible that the effect disappeared over time. This would be consistent with the hypothesis, because dependency on priming would be stronger the more innovative an expression is. But given that the effect in the 1980s is not

approach taken here is that some turns in Parliamentary debates are (by conversational standards) extremely long, making for a considerable distance between the assumed prime and its target.

⁶ It is easy to demonstrate that the data show a persistence effect. For example, in the Parliamentary debates from the 1980s, *key* occurs at an average rate of about 1 hit per 81 turns. But in the turns that immediately follow an occurrence of *key* that rate goes up to 1 hit per 26 turns. The main methodological leap made here is the assumption that this persistence effect can indeed be interpreted as a priming effect.

very strong to begin with, it may also simply be a fluke of the data. As such, the evidence here is still tentative. Second, if the hypothesized priming effect exists, it should be clear that it cannot be a major propagator of change. In the corpus, the primed condition is uncommon and the great majority of adjectival and debonded instances of *key* occur outside primed contexts.

4.2.2 Collocational priming

If collocational priming promotes innovations in the development of *key*, it is expected that new uses of *key* are more likely to surface when *key* has been primed by a contextually available lexical element. The prediction can be tested with some precision on the non-predicative adjectival and debonded uses of *key*. Since in these uses *key* is always a premodifying element, it always pattern with a nominal head that associates with the form *key* to a greater or lesser degree. For example, it is intuitively clear that *role* in (9a), being the head of a frequent *key*-compound (*key role*), more strongly associates with *key* than *land* in (9b).

- (9) a. That is a *key* and important role. (1988, HC)
b. The question of *key* industrial land bears directly on harbours. (1973, HC)

Intuitions about collocational association can be backed up by Mutual Information (MI) scores based on joint occurrence in the Hansard Corpus. As expected, *key role* is then found to be a strong collocation with a mutual information score of 7.6, whereas *key land* is simply unattested. This means that, other things being equal, *key* is more likely to have been collocationally primed in (9a) than in (9b).

Pursuing this logic, MI scores for *key* and its nominal head were calculated for all the attested debonded and adjectival uses of *key*. The attested instances were collected from the data used to analyze individual variation in Section 4.1 above. MI scores were calculated on the basis of the complete 1970-1999 portion of the Hansard Corpus.⁷ The results are visualized in Figure 7. It is found that adjectival uses of *key* tend more strongly to occur with established *key*-collocates than debonded uses. Paradoxically, then, the construction that is grammatically further removed from *key*-compounding is collocationally more like it. Yet, this is in line with the prediction. Grammatically, *key* is least likely to appear in purely adjectival contexts, but this is compensated for by the presence of a *key*-collocate. Debonded uses of *key*, which are grammatically less deviant and are already better established, depend less heavily on the presence of a *key*-collocate. Still, taking an MI of 3 or more as threshold for linguistic significance (Stubbs 1995), even debonded uses are built on well-established *key*-compounds

⁷ Unattested combinations, such as *key land*, have been assigned an MI score of 0. For instances of one-substitution, the head has been restored from the context. For example, in *Targets will also need to be set—I have outlined some key ones today* (1998, HC) the (restored) head is *targets*.

in about half of all cases (110 out of 211, i.e. 52.1%).⁸ All this fits in with the idea that innovations to an expression's use tend to occur under conditions that improve that expression's retrievability.

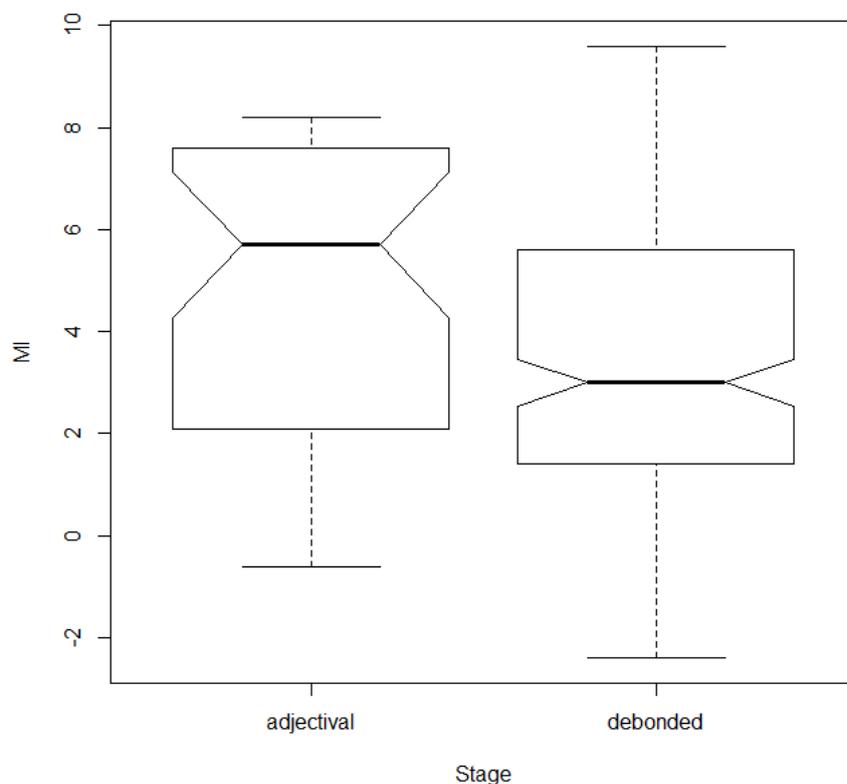


Figure 7. MI scores for *key* and its head in premodifying debonded and adjectival constructions in the Hansard Corpus.

5 Conclusions

The history of English *key* is a modest change, hardly complicated and with minimal impact on the structure of English. Moreover, the investigations into its underlying mechanisms, while making use of an exceptionally rich data source, still run up against the limits typical of a finite historical record. That said, and at the inevitable risk of overgeneralization, a theoretical point can be made. The evidence presented above is largely consistent with the predictions made by the cascade model. As the constructions resorting under an earlier stage of a change become better entrenched, they more easily spark off the innovations that constitute the following stage. This is visible in patterns of variation across individual speakers, as well as in the changes speakers undergo in the course of their lifetimes.

⁸ That this proportion is higher in the 1970s (61.1%) than in the 1980s (48.5%) and 1990s (51.4%) may suggest that debonded uses, too, used to be more dependent on existing *key*-collocations at earlier stages. However, the trend is not significant (χ^2 1.40, $p = 0.24$; $\phi = 0.08$, comparing the 1970s to the 1980s-1990s).

Similarly, when earlier-stage constructions have been primed, the chances for innovative constructions to appear improve. The priming effect is weakly evidenced in contexts involving direct priming from previous discourse, and it is seen, more convincingly, in contexts involving indirect priming through the contextual presence of collocates.

From these findings we may begin to glean a clearer picture of the language-internal factors that drive gradual change. Expressions expand their usage contexts along pathways determined by similarity relations, driven on by local analogies. Those chains of analogical extensions are triggered by the interaction between conventionalization (entrenchment) and linguistic creativity (innovation). Creativity and innovation are unlikely and costly, but given conventional, readily available source material to work from, it comes within the language user's reach. This way, it appears, expressions' usage can continually expand to the fringes of what is grammatically conceivable. And when the conceivable becomes fully acceptable and conventional, the edges of conceivability, in turn, expand outwards. What is there to stop this expansion? Perhaps processes of gradual change peter out when they fall short of nearby analogs ([AUTHOR & OTHER] *forthc.*). Possibly, the conventionalization of innovative patterns is the crucial bottleneck (Croft 2010), which may ultimately depend on an entirely different set of factors, such as functional need or extralinguistic motivations.

Acknowledgements

[ACKNOWLEDGEMENTS]

References

- Aitchison, Jean. 1991. *Language change. Progress or decay?* Cambridge: Cambridge University Press.
- Bergs, Alexander. 2005. *Social networks and historical sociolinguistics. Studies in morphosyntactic variation in the Paston Letters (1421-1503)*. Berlin: Mouton de Gruyter.
- Chappell, Hilary. 2008. Variation in the grammaticalization of complementizers from *verba dicendi* in Sinitic languages. *Linguistic Typology* 12: 45-98.
- Croft, William. 2010. The origins of grammaticalization in the verbalization of experience. *Linguistics* 48: 1-48.
- Denison, David. 1986. On word order in Old English. *Dutch Quarterly Review* 16: 277-295.
- Denison, David. 2001. Gradience and linguistic change. In Laurel J. Brinton (ed.). *Historical linguistics 1999*. Amsterdam: John Benjamins. 110-144.
- Denison, David. Forthcoming. *English word classes: Categories and their limits*. Cambridge: Cambridge University Press.

- Durrant, Philip, and Alice Doherty. 2010. Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*. 6: 125-155.
- Fischer, Olga. 2007. *Morphosyntactic change. Functional and formal perspectives*. Oxford: Oxford University Press.
- Güldemann, Tom. 2002. When 'say' is not say: The functional versatility of the Bantu quotative marker *ti* with special reference to Shona. In Tom Güldemann and Manfred von Roncador (eds.). *Reported discourse. A meeting ground for different linguistic domains*. Amsterdam: Benjamins. 253-287.
- Harris, Alice C. and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Haspelmath, Martin. 1998. Does grammaticalization need reanalysis? *Studies in Language* 22: 315-351.
- Hoey, Michael. 2005. *Lexical priming. A new theory of words and language*. London: Routledge.
- Langacker, Ronald W. 1977. Syntactic Reanalysis. In Charles N. Li (ed.). *Mechanisms of syntactic change*. Austin: University of Texas Press. 57-139.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar*. Vol. 1. *Theoretical prerequisites*. Stanford: Stanford University Press.
- Lee, Ji Won. 2011. Much ado about a lot: A corpus study of *much* as a negative polarity item. Paper presented at the 20th International Conference on Historical Linguistics in Osaka, Japan. July 26th, 2011. (Downloadable from <http://www.acsu.buffalo.edu/~jiwonlee/>.)
- Naro, Anthony J. 1981. The social and structural dimensions of a syntactic change. *Language* 57: 63-98.
- Norde, Muriël. 2009. *Degrammaticalization*. Oxford : Oxford University Press.
- Paunonen, Heikki. 1976. Idiolectal variation in Helsinki urban speech. *International Journal of the Sociology of Language* 10: 125-140.
- Petré, Peter. 2012. General productivity: How become waxed and wax became a copula. *Cognitive Linguistics* 23: 27-65.
- Petré, Peter, and Freek Van de Velde. 2014. Tracing real-life agents' individual progress in ongoing grammaticalization. How grammaticalization processes create grammar: from historical corpus data to agent-based models. Paper presented at *EvoLang*, Vienna, 14-17 July.
- Rosemeyer, Malte. 2014. *Auxiliary Selection in Spanish. Gradience, Gradualness, and Conservation*. Amsterdam: John Benjamins.
- Schmid, Hans-Jörg, and Annette Mantlik. Fortchoming. Entrenchment in historical corpora? Reconstructing dead authors' minds from their usage profiles. *Anglia*.
- Stubbs, Michael. 1995. Collocations and semantic profiles. On the cause of the trouble with quantitative methods. *Functions of Language* 2: 1-33.

- Szmrecsanyi, Benedikt. 2006. *Morphosyntactic persistence in spoken English. A corpus study at the intersection of variationist sociolinguistics, psycholinguistics, and discourse analysis*. Berlin: Mouton de Gruyter.
- Tabor, Whitney. 1994. *Syntactic innovation. A connectionist model*. Doctoral dissertation. Stanford: Stanford University.
- Timberlake, Alan. 1977. Reanalysis and actualization in syntactic change. In Charles N. Li (ed.). *Mechanisms of syntactic change*. Austin: University of Texas Press. 141-177.
- Vartiainen, Turo. 2013. Subjectivity, indefiniteness and semantic change. *English Language and Linguistics*. 17: 157-179.
- Weiner, E. Judith, and William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19: 29-58.