English Language and Linguistics (2025), 1–17
doi:10.1017/S1360674325000206

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Tied together, fried together: understanding semantic extensibility through semantic loss

Hilke Ceuppens and Hendrik De Smet

Department of Linguistics, KU Leuven, Blijde Inkomststraat 21, 3000 Leuven, Belgium
Corresponding author: Hilke Ceuppens; Email: hilke.ceuppens@kuleuven.be

(Received 6 November 2024; revised 14 March 2025; accepted 17 March 2025)

Abstract

Semantic extensibility captures the semantic side of productivity. It is the likelihood that a given sense of a linguistic expression will support extension to new senses. Even though linguistic expressions are naturally polysemous, semantic extensibility is constrained. In previous literature, it has been argued that semantic extensions are motivated by mostly one-directional conceptual operations such as metaphor and metonymy, and that in any polysemous expression only one or a few so-called 'sanctioning' senses have privileged status in supporting new extensions. One factor believed to determine sanctioning status is high frequency. Drawing on three case studies from the history of English, involving change in the adjective *awful*, the preposition and adverb *about* and the multifunctional item *so*, this article provides diachronic evidence from semantic loss to support this view. On the one hand, it is shown that when old sanctioning senses go into decline, this also impacts the senses derived from them, underscoring the motivational relations that tie extended senses to sanctioning senses. On the other hand, what typically initiates a decline in a sanctioning sense is a frequency increase elsewhere in the polysemy network coincident with the emergence of a new sanctioning sense, underscoring the role of frequency in determining sanctioning status and the directionality of sanctioning relations.

Keywords: frequency; polysemy; productivity; semantic extension; semantic loss

1. Introduction

Productivity can be thought of as a gradient property of any linguistic element that captures the likelihood language users will use it in new ways. Baroddal (2008) has labelled this property 'extensibility'. This extensibility is usually understood in formal terms and then applies to how freely forms distribute over other forms by entering into new formal pairings. Clearly, this is also what Bolinger had in mind when he defined productivity as 'the statistically determinable readiness with which an element enters into new combinations' (1948: 18). For example, English weak past tense formation has high extensibility, because the past tense *-ed*-suffix will automatically extend to attach to any verb that is newly introduced into the language. All French-origin verbs that entered English in the fourteenth century form their past tense by *-ed*-suffixation (*appealed*, *cried*, *retained* and so on), just as any new verb stem adopts the weak past tense in PDE (*booped*, *cancanned*, *windbagged*). In contrast, extensibility is low for abstract noun formation through *-th* (*depth*, *warmth* but

thickth, *calmth*). We will refer to this dimension of productivity here as combinatory extensibility.

In the current literature on productivity, semantics largely plays second fiddle. Its relevance is recognized but only in as far as it bears on combinatory extensibility. On the one hand, it is generally accepted that combinatory extensibility correlates with semantic transparency. A form will be productive only in as far as its meaning remains predictable whichever combination it enters into. Barðdal (2008) argues that semantic coherence is a predictor of a construction's productivity. Hay (2003) has demonstrated that low productivity correlates with low semantic transparency. On the other hand, it makes sense that combinatory extensibility is constrained by the semantic compatibility between elements (Goldberg 1995). For example, the prefix *un-* can combine with verbs to denote a process involving a reversal of the process denoted by the verbal stem. That is why almost all verbal stems *un-* combines with express accomplishments and achievements (e.g. *undress*, *unload*, *unlock*, *untie*). These verbal stems prototypically denote a process that involves both a change of state and an inherent endpoint, making their reversal easily conceivable. Attached to verb stems that prototypically express states or activities, *un-* becomes awkward or ungrammatical (e.g. **undance*, **unstand*), or else it coerces an accomplishment reading (e.g. *unknow*).

However, our conception of productivity can embrace the semantic dimension more fully. It is obvious that just as linguistic elements may or may not extend in combinatorial space, they may or may not extend in semantic space. A conceptual unification[1] of productivity along the different dimensions of linguistic organization has been proposed by Langacker (1987) through his concept of 'sanction'. Langacker defines sanction as a correspondence between a conventional unit and a target item, which leads to the unit sanctioning the target (1987: 65–71). Full sanction describes the situation where conventional unit and target item are fully compatible and corresponds to the linguistic notion of well-formedness. Partial sanction describes the situation where the correspondence is (as yet) unconventional and involves extension. The following example illustrates the difference between full and partial sanction:

> Consider a child in the process of learning the various senses (conventional usages) of the word *tree*, and suppose the term is first applied, in his early experience, to such ordinary specimens as oaks, maples, and elms. Their perceptual prominence and obvious gross similarities enable the child to extract a conception that embodies their commonality, while excluding the many properties that vary from one instance to the next. ... With continued usage, this initial conception becomes more deeply entrenched, and comes to be invoked for the categorization of more divergent experience. ... Suppose, then, that our child has mastered the concept [TREE] (the eventual category prototype), as well as the symbolic relationship [[TREE]/[tree]]. When he encounters a tall plant with branches, leaves, and bark he readily sees it as conforming to the specifications of [TREE] and takes it as a straightforward instance of the *tree* category. What happens, now, when he first encounters a pine, which is [TREE]-like in most respects but has needles instead of leaves? He will quickly learn to call it a *tree*, either from hearing someone refer to it in this way or because this is the most nearly applicable term at his disposal. This usage implies the symbolic unit [[PINE]/[tree]], derived by extension from the original [[TREE]/[tree]]. (Langacker 1987: 373–4)

[1] Within the narrower domain of grammaticalization, Himmelmann (2004) unifies the various dimensions of productivity under his three different types of 'expansion'. His host-class expansion roughly corresponds to our combinatory extension, and his semantic-pragmatic expansion subsumes our semantic extension.

But exactly what determines extensibility? On a usage-based perspective, extensibility must arise from the state of speakers' internalized language system at any given time. In the above quotation, Langacker links both schema formation ('to extract a conception that embodies their commonality') and entrenchment ('continued usage') to extensibility ('categorization of more divergent experience'). In this spirit, De Smet suggests that 'easy availability of a strongly entrenched symbolic relation facilitates its over-application through partial sanction. The result must be semantic change' (2017: 80). This certainly aligns well with the current understanding of combinatorial extensibility. In the domain of morphology in particular there is broad consensus about the factors promoting or reducing extensibility. Among those are the semantic constraints mentioned above but also, and more importantly, the complex interplay between type and token frequencies that determines how any given combinatory pattern becomes both schematized and entrenched (Baayen & Lieber 1991; Plag 1999; Hay & Baayen 2002; Goldberg 2006; Bybee 2007; Barðdal 2008; Schmid 2018).

Where it comes to semantic extensibility, the literature advances detailed proposals. Specifically, in work on polysemy it has been argued that word senses cluster into radial networks, with the majority of senses deriving from a single core sense or what Tyler & Evans (2003) (in an echo of Langacker 1987) call the 'sanctioning sense'. This view makes predictions about where the potential for semantic extension lies for a given linguistic element, and how semantic extensibility is constrained – more on which below. At the same time, acceptance and consensus in the field have proven elusive. Radial networks raise a range of practical as well as theoretical questions that have so far not been satisfactorily addressed. While these issues should justify a certain reluctance to uncritically embrace Radial Network Theory, it would be a mistake to throw out the child with the bathwater. This article seeks to revisit some of the ideas from Radial Network Theory on what determines the potential for semantic extensibility and explores their applicability to diachronic semantic change. To this end, evidence is drawn from the counterpart to semantic extensibility, namely semantic loss. It is argued that cases of semantic loss support the kinds of sense clusterings and directional sanctioning relations posited by Radial Network Theory, and moreover point to the role of frequency as one of the factors that can elevate specific senses to sanctioning status.

In what follows, we start by discussing Radial Network Theory and its contribution to our understanding of semantic extensibility, but also its current shortcomings (section 2). Next, we discuss a number of specific changes, each of which involves the (near-)loss of a historical sanctioning sense in a linguistic element (section 3). The expressions studied range across the lexical-to-grammatical continuum. They are the adjective *awful*, the preposition and adverb *about*, and the multifunctional element *so.* Two patterns come to the fore. First, in each case the loss of a historical sanctioning sense coincides with a marked frequency increase in one or more other senses, likely reflecting the emergence of a new sanctioning sense. Second, loss of an old sanctioning sense typically extends also to other senses that derive directly from it. When an old sanctioning sense obsolesces, this also implicates the other senses that are derived from it and that receive no sanction from the newly emerging sanctioning sense. These findings support some of the basic intuitions underlying Radial Network Theory. That is, senses stand in directional motivational relations to one another, and frequency is one determinant of the likelihood that a sense can form the basis for semantic extensions (section 4).

2. Radial networks and semantic extensibility

Radial Network Theory (RNT) subsumes a range of closely related proposals that seek to describe and explain polysemy by laying out a number of principles of semantic organization. Its roots lie in the Cognitive Linguistics work of the late twentieth and early twenty-first

century (Brugman 1981; Lakoff 1987; Brugman & Lakoff 1988; Tyler & Evans 2001, 2003; Evans 2004, 2005). RNT maintains that senses stand in motivational relations to one another, in that they are linked through basic conceptual operations such as metaphor or metonymy. Because senses are connected by these motivational ties, they can be thought of as organizing into networks that reflect synchronic (and diachronic) relations of derivation. RNT also holds that at least one of the senses in such a network will have privileged status as the core or sanctioning sense that gives (or gave) rise to the majority of derived senses. This is why networks have the radial structure that gives them their name.

As a toy example, consider the noun *door*. A plausible sanctioning sense for *door* would be 'movable barrier serving to close or open a passage into a room' (*OED* 1a), as in (1a), which gives rise to derived senses including 'opportunity' (*OED* 3) through metaphor as in (1b) or 'opening or passage into a room' (*OED* 2) through part–whole metonymy as in (1c). Figure 1 visualizes these sense relations as a radial network. Finally, motivational relations between senses are mostly directional – for example, in metaphorical extensions it is typically (if not exclusively, see Allan 2021) the more concrete embodied senses that support extension to more abstract senses, rather than vice versa.

(1) (a) So as I'm still trying to stuff things away, she opened the **door** and everything fell out. (1991, BNC)
(b) His madness is the **door** to sanity. (2009, COCA)
(c) He stood in the **door** of the milking house, holding out the buckets for her to take. (1989, BNC)

This model can in principle explain why some semantic extensions may gain or lose in likelihood over time. When a new extension gains in prominence in a network, it can assume core status and begin sanctioning further extensions. Conversely, because sanctioning relations are mostly one-directional, it will not support the old sanctioning sense, nor any of the senses derived from the old sanctioning sense. Thus, over time, a peripheral extension in the network may become more salient, to the point that the prototypical centre shifts from the old core to the extension. Györi (2002) gives the example of PIE **(s)keu* 'to cover' to Germanic *huson* 'covering for the legs' and eventually English *hose*, where the extension 'covering the legs' becomes culturally salient and prototypicalizes at the expense of the older centre. At this point, the new prototype can start sanctioning new extensions like the metaphorical 'flexible tube'. According to Györi, the process is supported by the extension's original low degree of prototypicality, which leaves it less strongly integrated in the network and makes it easier for it to detach as a new prototype. Baumann *et al.* (2023) call attention to the role of frequency in the shift.

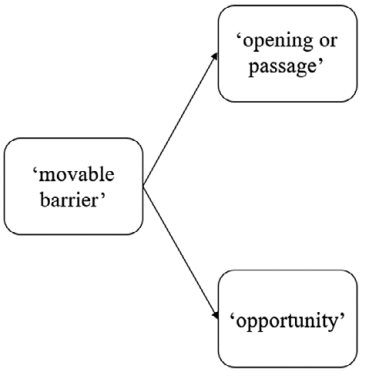

Figure 1. Toy radial network for *door*

While RNT could potentially explain part of the dynamics of semantic loss, there is no denying that RNT faces major challenges. First, with its emphasis on a core and periphery, the model clearly draws inspiration from Prototype Theory (based on Rosch 1978; for influential work see i.a. Geeraerts 1997 or Taylor 1995), but in doing so it more or less tacitly extends the idea of prototypical organization from purely conceptual categories to the larger constellations of conceptual categories that make up the semantic pole of linguistic expressions. As far as conceptual categories go, the evidence for prototypical organization is strong, but to assume that the same organizational principle can be scaled up to describe polysemy is arguably a leap of faith (however, see Cai & De Smet 2024a: 4).

Second, RNT struggles to disentangle synchrony from diachrony. Patterns of diachronic extension offer some of the best evidence of motivational relations between senses. But semantic extensions that were once motivated may no longer be so in synchrony. For example, the extension in *film* from 'thin skin or layer' to 'movie' only made sense as long as movie production typically involved the use of thin sheets with a layer of photosensitive chemicals. In other words, synchronically, not all senses of an expression will be equally motivated, but little effort has been made to further theorize the role of connection strength in radial networks. Recent psycholinguistic studies have additionally shown that metaphorical and metonymic extensions are processed differently (Lopukhina *et al.* 2018), underscoring the idea that not all senses and sense connections are equal.

Third, and most problematically, RNT lacks both a unified representational format and a well-circumscribed method to identify senses and sense relations (Sandra & Rice 1995). Whereas Lakoff (1987) adopts a fine-grained full-specification approach whereby every sense in the network has its own image schema, Tyler & Evans (2003) argue for a principled polysemy model, which attempts to constrain the number of senses and place an increased importance on the role of real-world knowledge and context by distinguishing between full senses and vagueness. The tests they propose do not, however, fully remedy the problems of classical polysemy tests (Geeraerts 2015). Lumping/splitting controversies inevitably persist.

In the same way, RNT faces problems determining the sanctioning sense in a radial network (Gilquin & McMichael 2018). A variety of criteria have been suggested to solve this problem. For instance, the sanctioning sense in a radial network is said to be experientially more basic, more frequent, first listed under elicitation, older, the basis for morphological derivations and (perhaps somewhat circularly) the most plausible basis from which to derive other senses. But these various criteria may be at odds or simply hard to apply in practice. Moreover, even when one sense supports a majority of extensions, other senses will often support further extensions, at least suggesting the possibility that a single network may sustain multiple cores (Langacker 1987: 396).

New approaches to semantic representation have primarily addressed the methodological issue of sense discrimination by leveraging distributional evidence to analyse polysemy (Gries 2010; Heylen *et al.* 2015; Hilpert & Correia Saavedra 2017; Fonteyn 2018; Budts 2020). The family of distributional methods has certainly infused a healthy dose of objectivity into the debate, and has proven very apt at NLP tasks. On a theoretical level, its successes have encouraged a view of meaning as a contextually determined point in a continuous and multidimensional semantic space, challenging lexicographic orthodoxy and the very notion of senses (Li 2024). Unlike RNT, however, this type of approach says little about the constraints on semantic extensibility.

For this reason – and despite the very real problems with the model – some of the basic intuitions underlying RNT continue to be worth exploring, as they offer potential insight into the dynamics of semantic change. Specifically, this article builds on the concepts of sense clustering, extensibility and directionality. First, senses cluster into networks structured by the motivational ties that connect them. In such networks some senses are more closely tied to each other than others, leading to groupings of related senses. Second, some

senses in these networks have privileged status as sanctioning senses in that they provide the motivation that both supports new semantic extensions and that sustains existing polysemy. For example, Petré frames the core use of a multifunctional expression as 'a type of glue by which all uses can stick together' (2014: 74). Similarly, Huang *et al.* (2021) show that the more prototypical senses of Chinese 火 *huǒ* 'fire' give rise to more extensions. They cast this in terms of productivity:

> the more typical and salient senses of the FIRE character seem to be more productive in terms of the number of the extensions which develop from them. For instance, there are more extensions that develop from LIGHT (three extensions) than from FLAMES (only one extension). (2021: 38)

The idea that polysemous networks consist of more central (core) senses which are productive and sanction peripheral uses is also invoked by Cai & De Smet (2024b: 57), who state that:

> Sanctioning can be thought of as a form of semantic productivity: the more salient and core-like the sanctioning sense, the stronger the support for any extensions based on it and the more likely the use of any of the resultant secondary senses becomes.

This ties in with the idea that innovations to an expression are likely only to the extent that they are supported by well-entrenched uses (De Smet 2016: 86). The frequency of a sense, then, should be one of the determinants of its extensibility. Third, sanctioning is typically directional, in that the core support its extensions but not vice versa. This is well-supported by the observation that, in diachrony, recurrently attested sense extensions are not reversible (e.g. 'door' will give rise to 'opportunity' but not vice versa). Therefore, it is plausible that in synchrony, too, a sanctioning sense supports its extensions, but the extensions do not necessarily support the sanctioning sense. When sanctioning status shifts to a new sense, the new sanctioning sense may no longer sustain the old sanctioning sense from which it once derived, resulting in the loss of the old sanctioning sense.

3. Case studies

In what follows, we examine both causes and consequences of diachronic shifts in what constitutes the likely sanctioning sense in the radial network of a polysemous expression. The discussion draws examples across the lexicon-grammar continuum with *awful* (section 3.1), *about* (section 3.2) and *so* (section 3.3).

As a methodological note, the semantic analyses presented below are based on the semantic analyses in the *Oxford English Dictionary* (*OED*). Although this article revisits concepts from RNT, it does not seek to adopt all its ideas, let alone solve its shortcomings. These include the difficulties of deciding whether senses should be split or lumped together. By drawing on the semantic descriptions in existing *OED* entries we have relied on a properly independent source to ensure the best level of replicability for our analyses, and as a guardrail against confirmation bias. It should be noted that by its nature as a reference work, the *OED* leans towards a splitting approach to semantic description, often drawing maximally subtle sense distinctions. Where we deviate from its analyses it has typically been to lump senses together, so as to keep corpus annotation doable.[2]

[2] Full annotated samples for all case studies are available online from https://doi.org/10.48804/KNZBCL

3.1 Awful

The dominant sense of *awful* today, 'unpleasant, bad', originated in colloquial English at the end of the eighteenth century. This first case study evaluates the impact of the new extension on the older senses and further development of the adjective in terms of extensibility. It does so by establishing the radial network of the adjective on the basis of the relevant *OED* entry and corpus data, and by mapping the frequency changes of its different senses from Early Modern English to Present-Day English.

The *OED* entry for *awful* lists nine senses. Three of those senses reflect the use of *awful* as an intensifier (emphasizing extent, something pleasant, and something enjoyable, respectively I.3.a, I.3.b and I.3.C in the *OED* 3 entry) and have been collapsed into one category in what follows. The sense 'frightened or apprehensive of doing something' is not further discussed in this article as it does not occur in the corpus data that were analysed (see below). In this way, eight senses can be distinguished for discussion.

The historically earliest sense of *awful* is 'inspiring fear' (2a), straightforwardly reflecting the meaning of its stem *awe* 'reverential fear, admiration'. This oldest sense sanctions the use of secondary, related senses, specifically 'arousing respect' (2b) and 'filled with fear' (2c). The former differs from 'inspiring fear' in its positive connotation, which is often signalled by other coordinated adjectives. It next forms the basis for a further extension, 'demonstrating respect' (2d). These senses make up the semantic profile of *awful* until the beginning of the nineteenth century, when a new sense extends from 'inspiring fear'.

(2) (a) The nations of antiquity, careless of each other's safety, were separately vanquished and enslaved by the Romans. This **awful** lesson might have instructed the Barbarians of the West to oppose, with timely counsels and confederate arms, the unbounded ambition of Justinian. (1776, CLMET)
(b) any of the robust trees of the forest which we consider as beautiful; they are **awful** and majestic; they inspire a sort of reverence. (1757, CLMET)
(c) And now all was **awful** expectation in the court, and every heart was full of grief for Antonio. (1807, CLMET)
(d) we are willing to bow to the splendid equipage, and stand at an **awful** distance from the pomp of a princely estate (1767, CLMET)

The new extension, 'unpleasant' (3a), retains the negative connotation of 'inspiring fear' but is weaker and more general in use. The situations or persons it refers to do not evoke an emotion as strong as fear, but a general disdain. As this sense gains in prominence, it starts sanctioning the new uses 'of poor quality' (3b) and 'unwell, troubled' (3c). Second, an intensifier use (3d) develops from 'inspiring fear' and 'unpleasant', emphasizing an unpleasant or conversely enjoyable experience, or a large amount. This use is most apparent in the collocation *an awful lot.*

(3) (a) Yeah I was really attracted to him but I just could not speak to him, it was **awful** (1985–94, BNC)
(b) Early PC-based software was both expensive and fairly **awful** (1985–94, BNC)
(c) No you say that I once got off with someone for some cigarettes, I never felt so **awful** (1985–94, BNC)
(d) She did Rome in a swift two days, gave half the time to Venice, but vows that she saw everything, although in **awful** haste (1890, CLMET).

The eight senses can be regrouped into three clusters on the basis of their diachronic connections on the one hand and their semantic relatedness on the other. A first group consists of the oldest senses, (2a) to (2d). Later extensions (3a) to (3c) with a weaker

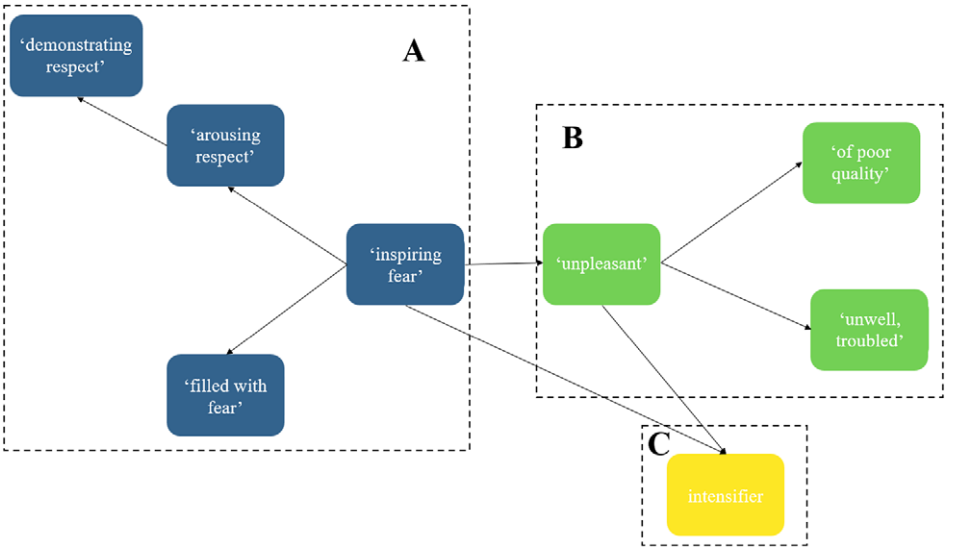

Figure 2. Radial network for *awful*

expression of emotion make up the second cluster. Last, (3d) differs from all other senses in activating the scalarity in the element it modifies. Figure 2 visualizes the semantic development of *awful.* The successive clusters are enclosed in dashed lines and labelled chronologically (A representing the oldest cluster, C the most recent). The arrows between senses show the direction of the extensions.

To map the frequency changes of the (clusters of) senses of *awful*, the *Corpus of Late Modern English Texts* (CLMET, version 3.1) and the written sections of the *British National Corpus* (BNC Consortium 2007) were queried for all spelling variants of the adjective. CLMET is divided into three subsequent periods of 70 years, namely 1710–80, 1780–1850 and 1850–1920. The BNC covers the last two decades of the twentieth century.[3] Per period, a random sample of 200 hits was annotated for sense and sense cluster. Ambiguous instances (most often between an intensifier use and another sense) were labelled as such. Though somewhat more frequent in the third period, they are generally few and are not further discussed in what follows.

Figure 3 shows that all attested senses in the first period belong to the A-cluster. *Awful* is used in the sense 'unpleasant' in almost a third of all cases, but the other senses, 'arousing respect', 'filled with fear' and 'demonstrating respect', all have a substantial share in the semantic make-up of *awful.* The second period is likewise characterized by the dominance of the A-cluster. The first sense of the newer B-cluster, 'unpleasant', and the intensifier use occur but remain infrequent, while'of poor quality' occurs only once and 'unwell' is not yet attested. The second half of the nineteenth century marks a notable shift in the semantic profile of *awful.* The extension furthest removed from 'inspiring fear' in the A-cluster, 'demonstrating respect', is only attested once and 'filled with fear' has become rare, while the share of 'inspiring fear' has also declined. The share of 'unpleasant' conversely increases, as does the intensifier use. The sense 'unwell' is attested for the first time. By the end of the twentieth century, all senses of A but the oldest have disappeared (save for a single occurrence of 'arousing respect') and 'inspiring fear' appears to be on its way out. The

[3] A limited number of exceptions date back to before 1980.

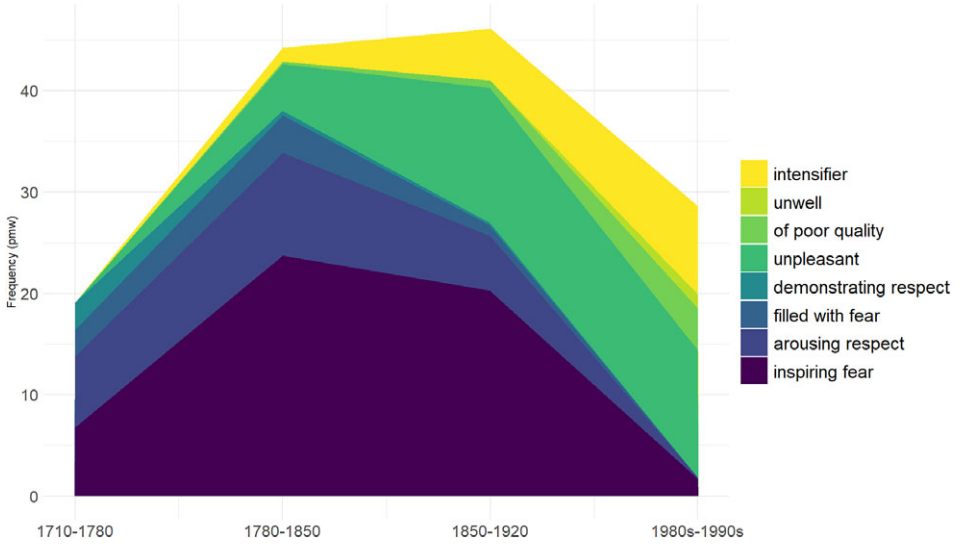

Figure 3. Semantic development of *awful* from Late Modern English (CLMET) to Present-Day English (BNC)

new dominant sense of *awful* is 'unpleasant' and the other B senses have increased in frequency. The intensifier use likewise continues to increase.

Based on the frequency data and the potential for licensing new extensions, 'inspiring fear' and 'unpleasant' are the clear successive cores of the network. The other senses in A and B form the periphery, with 'demonstrating respect' as the most peripheral sense. When *awful* extends to 'unpleasant' and this extension takes off, the share of the older senses begins to decline. The first sense to be lost is the furthest extension into the periphery, followed by the other peripheral senses. At this point, the old sanctioning sense begins to lose its potential to sustain its various extensions. For its part, the new sanctioning sense begins to support new extensions, forming the emerging B-cluster, but it fails to support the old sanctioning sense, which also goes into decline. By the beginning of the twenty-first century, the old sanctioning sense is the only sense of the A-cluster to survive, albeit at a very low frequency

3.2 About

Just as for *awful*, the *OED* was used to establish a semantic network for *about.* The *OED* distinguishes four major clusters of senses. The first cluster, A, henceforth called POSITION, contains all static position senses of *about.* This includes the oldest sense of *about*, 'encircling' as in (4a). This sense gives rise to 'accompanied by', 'in every direction' and 'in the vicinity of'. The latter, exemplified in (4b) later extends to 'near, within reach' as in *have your wallet about you*, and its figurative counterpart as in *have your wits about you.* A new dynamic sense, 'moving in a circuit around', also arises from the static 'encircling'. Together with 'moving in various directions' (4c), it makes up the B-cluster, MOTION. Second, from the A-cluster (and specifically the 'near' sense), an approximator use of *about* extends. This includes approximation in time as in (4d), approximation in number and the general approximator 'approximately, more or less'. Note that *about* is an adverbial preposition in (4c) and a degree adverb in (4d).

(4) (a) his sonne Balthazers after him, was commaunded to bee cloathed with purple, and to put a chaine of golde **about** his necke (1602, EEBO)

(b) There was (saith he) a young man **about** Wittemberge, who beeing kept bare and needie by his father, was tempted by way of sorcerie (1590, EEBO)
(c) as to his jumping **about** for joy, in talking of me, it might be partly from the vivacity of his constitution (1779, CLMET)
(d) After the troubles had continued **about** three years, Antiochus, one of the officers of the household, executed with success the Imperial commission of restoring Chosroes (1776, CLMET)

The motion-cluster gives rise to a mental connection use of *about*, 'concerning, regarding' (5a), with a transfer from motion on a surface to mental motion. This sense later extends to a mental state sense, which describes emotions (such as *happy about*) rather than thought processes. Several senses extend from both 'concerning' and a second sense. For example, 'consisting of' (5b) connects to both 'concerning' and the figurative reach use described above. Example (5c) contains the sense 'occupied with', which is attached to 'concerning' and 'moving in various directions'. It later develops a more grammatical extension 'on the point of' as *in I'm about to leave.* All these senses make up the last cluster of senses, labelled MENTAL CONNECTION. Figure 4 summarizes the description of the senses of *about* in the form of a radial network.

(5) (a) She had already taken leave of him once, in the house, and heard all **about** the balloon and the sailor-aeronaut and the preparations (1908, CLMET)
(b) Life is **about** feeling (1989, BNC)
(c) Once or twice before I had had to send people **about** their business. (1905, CLMET)

Next we traced the frequencies of the senses using data from *Early English Books Online* (EEBO), CLMET and BNC. The first and third period from CLMET were selected, and data from EEBO were selected for two seventy-year periods, 1470–1540 and 1570–1640, based on general frequency changes of *about* in these periods. Spoken data from the BNC were excluded, to improve diachronic comparability. The corpora were queried for all spellings

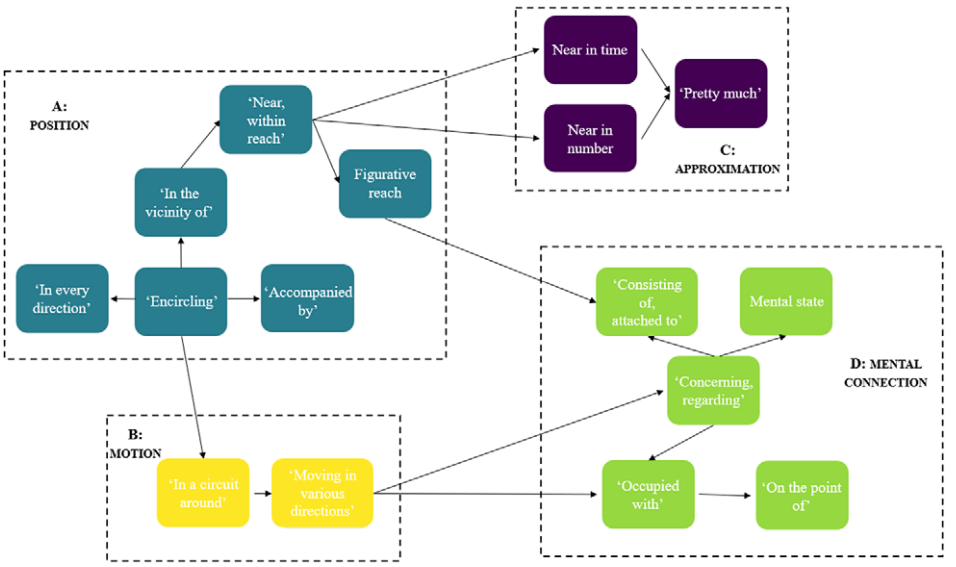

Figure 4. Radial network for *about*

of *about* and random samples were drawn of 200 hits per period. Annotation for sense and sense cluster was done according to the classification outlined above. Ambiguous or unclear instances were again labelled separately but are not further discussed below as they are very infrequent.

The frequency of the individual senses shows that 'encircling' (the source sense of *about*), approximation in time and number, and 'concerning, regarding' are the most common senses of *about*, respectively in the first, second and third and the last two periods. Because 'encircling' and 'concerning, regarding' are additionally central to the network as hubs for further extensions, they were identified as the core senses. For further analysis, clusters A and D were therefore split up into A1 and D1 for core and A2 and D2 for periphery (i.e. all other senses in the cluster). Figure 5 visualizes the normalized frequencies of the (sub-) clusters A1, A2, B, C, D1 and D2 throughout time.

Before the second half of the sixteenth century, the oldest cluster is the most frequent. All peripheral senses together outnumber the core, but individually none of them are ever nearly as frequent as 'encircling'. After this first period the frequency of the core and its periphery begins to steadily decline to the point of near-disappearance. By the end of the twentieth century, POSITION does not feature in the data except for two occurrences of 'in the vicinity of'. The motion cluster first slightly increases and later declines in frequency again. The approximation cluster likewise first increases in frequency, although the increase is steeper, and then stabilizes .

The core of the mental connection cluster steadily increases from the second period onwards. Its periphery behaves unexpectedly in the first period, when it actually outnumbers the core. There is an abrupt rise in frequency of the core after 1850, which is when the periphery really takes off in frequency as well. This rise of the mental connection senses causes an overall jump in the frequency of *about.* In the last period, cluster D, and especially its core, is dominant.

Cluster A starts out as the dominant one. Both this core and its periphery then begin to lose ground. Only 'in the vicinity of' outlasts the core, but with only two occurrences in the last period. By the second half of the nineteenth century, 'concerning' becomes much more

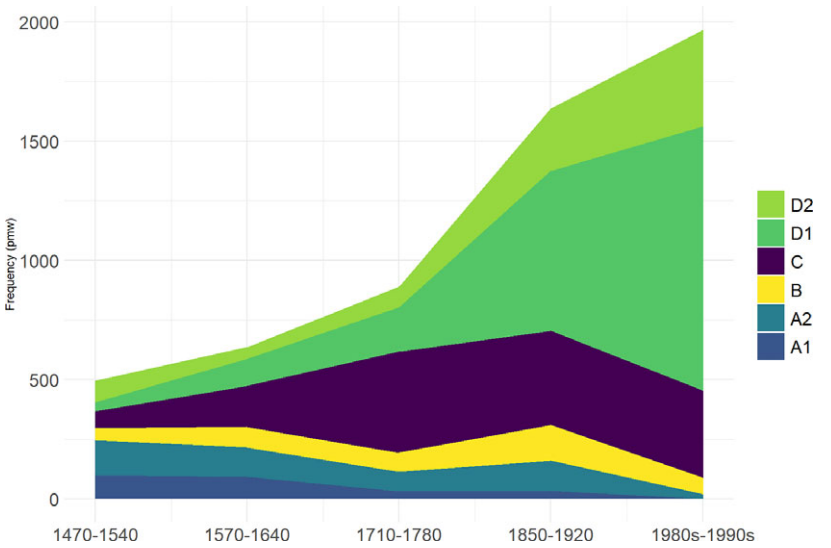

Figure 5. Semantic development of *about* throughout Early Modern English (EEBO), Late Modern English (CLMET) and Present-Day English (BNC) in terms of normalised frequency (per one million words)

frequent. At the same time, its periphery starts growing. At this point, the new core seems to be strong enough to start sanctioning new uses, but it does not support the old senses. The overall frequency of *about* is also impacted by this shift in core, as it more than doubles. The anomaly of the periphery outnumbering the core in the first period is the result of 'occupied with' being frequent. Arguably, this sense was already sanctioned by 'moving in various directions'. The motion and approximation clusters behave independently from the other clusters. They both first rise in frequency and then respectively slightly decline and stabilize by the last period. They are not affected in the same way peripheral senses are by the shift in core. It should be noted, however, that the frequency of the approximator use might be distorted by the vastly increased use of numbers in texts. This increase leads to more opportunities for using adnumeral markers and perhaps also greater need. In addition, the relative independence of the motion and approximation clusters may reflect their distinct syntactic status as adverb(-like) uses of *about.* This suggests that syntactic divides may to some extent insulate against the effects of semantic reorganization in the radial network.

3.3 So

The semantic development of *so*, rather than showcasing the dynamics of core and periphery, supports the idea of 'semantic glue' holding senses together. *So* appears in many different uses, ranging from manner adverb to discourse marker. Across this range of uses, its syntactic status is so diverse that connections across the whole network are probably weak at best, and a single synchronic core cannot be easily identified. However, uses that are plausibly related in meaning are seen to undergo the same fate over time, testifying to the role of local clusters in radial networks.

The manner adverb use is the oldest use of *so*, and has a deictic function (6a). It can be paraphrased as 'like this, in this way' and refers to the extralinguistic context. Similarly used deictically, but referring to a stretch of preceding (or occasionally following) discourse is place holder *so.* In (6b), *so* anaphorically stands for *really very fond of her.*

(6) (a) Nay, do not strut about **so** (1781, CLMET)
(b) For he is really very fond of her, and I hope will continue **so**. (1741, CLMET)

The second function of *so* is intensification. Intensifier *so* (7a) plausibly derives from manner and place holder *so.* The former is likely where *so* would have accompanied extraverbal comparison (e.g. through gesture), the latter where it would have involved comparison to the same property in a previously mentioned referent. Correlative *so*, which anticipates a following subordinate clause, is closely related to the intensifier use and often also functions as a booster (7b).

(7) (a) He is such a charming man, that it is quite a pity he should be **so** grave and so dull. (1811, CLMET)
(b) In the first place, the waves of the atmosphere cannot be **so** dangerous as those of the ocean (1821, CLMET)

Correlative *so* can also be used deictically (8a). Often, however, the meaning of correlative *so* is neither deictic nor solely intensifying. In most cases, the meaning of the following clause is resultative, working in tandem with intensifier *so*: the result functions as a precise measure of the intended degree of intensification. Example (8b) illustrates this mix of intensification and causality.

(8) (a) A feeling of shame and remorse took possession of William Dobbin as the broken old man **so** received and addressed him, as if he himself had been somehow guilty of the misfortunes which had brought Sedley so low. (1843, CLMET)
(b) He now seated himself by her, and talked **so** agreeably of Kent and Hertfordshire, of travelling and staying at home, of new books and music, that Elizabeth had never been half so well entertained in that room before (1813, CLMET)

The causal meaning is the sole meaning of conjunct *so* (9a). This form of *so* can be paraphrased as 'therefore' and establishes a causal connection to the previous clause. Subordinator *so* (9b) also has a causal meaning, introducing a subordinate clause of purpose or result. The subordinate clause is introduced by *that*, or occasionally *as.* The causal element is also important for discourse marker *so* (9c). It mimics a link to previous discourse, as if the utterance somehow logically follows from what was said before, but in fact preceding discourse may be lacking, and *so* essentially signals that the utterance is the speaker's logical response to the (extralinguistic) situation at hand. In (9c), a TV presenter comments on a climbing competition. As the competition has come to a close, the presenter signals a change in topic (which logically results from the competition having ended, but this is never made explicit).

(9) (a) there would be no crowding or confusion, for many people had gone away to the seaside, and **so** she was delighted at the thought of the picnic (1893, CLMET)
(b) Repeat what you think has been said **so** that any misunderstandings can be corrected (1992, BNC)
(c) Allan says it typifies what's good about British climbing open walls... Claire says it's always a challenge... you sometimes get cold... but she does it for enjoyment there's no other reason... **so** while our climbers rest and enjoy the sea air we can now catch up with the rest of the sporting action (1985–1994, BNC)

In summary, *so* can be used as a manner adverb or place holder with a deictic meaning, as an intensifier which functions as a booster, or as a conjunct, subordinator or discourse

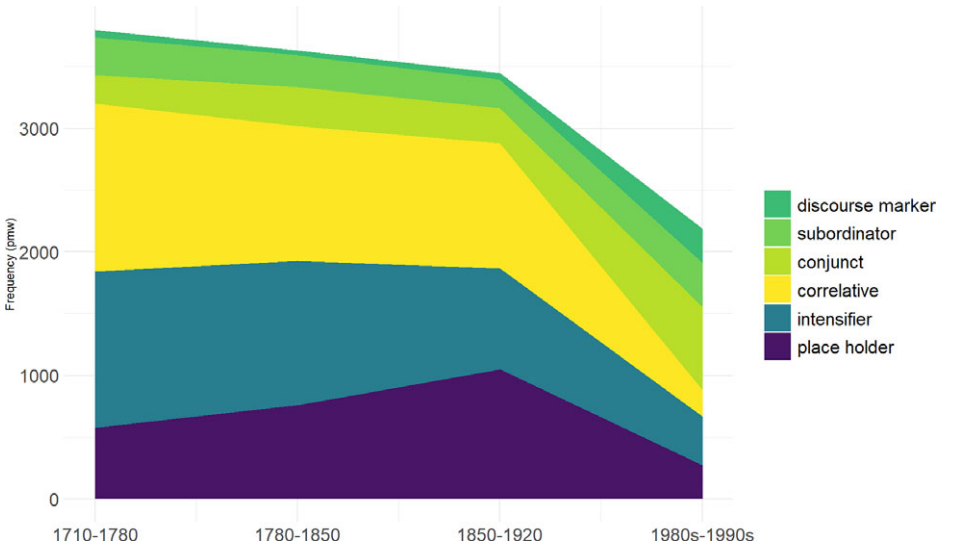

Figure 6. Development of the forms of *so* from Late Modern English (CLMET) to Present-Day English (BNC) in terms of normalised frequency (per one million words)

marker with a sense of causality. Correlative *so* appears with all three senses. To chart the development of *so*, CLMET and BNC were queried for *so* and random samples of 200 hits per period were annotated for form and function. The results are shown in Figure 6.

Manner adverb *so* occurs only once, in the second period, and it is therefore not included in the chart. The picture of *so* in the first two periods, until 1850, is one of stability. The intensifier and correlative are by far the most frequent. Place holder *so* is the third most frequent use and gains modestly in frequency in the third period. By the 1980s, however, it loses ground again, alongside the intensifiers and correlatives, which show a marked dip in frequency. This dip goes hand in hand with a general decline in frequency. On the other hand, the subordinator, conjunct and discourse marker use are on the rise. The frequency of the latter two respectively more than doubles and quintuples. *So* most frequently occurs as a conjunct in this last period, and the three categories expressing causality together far outnumber all other uses. A reversal of the situation at the start of the eighteenth century presents itself in terms of frequency: the uses of *so* with a deictic and intensifier sense have faded while the causal senses have become central.

A major reshuffling takes place in the use of *so* between the second half of the nineteenth century and the end of the twentieth Figure 6 repeats the pattern whereby decreases in one usage area of a polysemous expression are offset by increases in another. In addition, the changes directly reflect the major usage clusters of *so*. On the one hand, uses tied to the deictic and intensifier meanings decline whereas uses tied to causal meanings are stable or increase. What this shows is that the fates of different uses of *so* are tied together by their degree of semantic commonality. Supporting evidence comes from the break-down of change in the correlative uses in Figure 7 below. As discussed earlier, correlatives can express deictic meaning (see (8a) above), intensifying meaning (see (7b) above), or a combination of intensifying and causal meaning (see (8b) above). In our PDE data, excepting one instance of deixis, it is only the latter type that survives reflecting both the general weakening of deictic and intensifying uses of *so*, and the relative strength of uses that link to causal *so*.

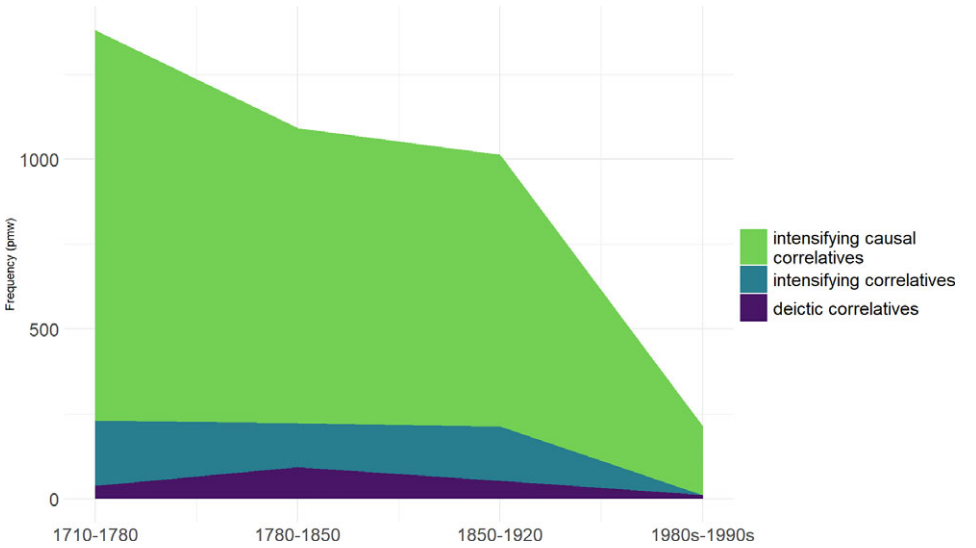

Figure 7. Semantic development of correlative *so* from Late Modern English (CLMET) to Present-Day English (BNC) in terms of normalised frequency (per one million words)

4. Concluding remarks

The case studies presented above support two related aspects of Radial Network Theory (RNT). First, they point to the role of frequency in determining which senses are more or less likely to support semantic extensibility. This does not mean that frequency is the only determinant of sanctioning status, but it is a plausible contributor. Second, shifts in what makes up the productive core of a radial network are found to coincide with shifts in the frequency profiles of different senses, and where an old core is seen to go into decline this is typically related to a parallel frequency increase for another sense or sense cluster. As sanctioning relations are directional, a new core will start sanctioning its own periphery, but it will not support the older senses. Third, patterns of semantic loss largely reflect the kinds of sense clusterings posited by RNT in that senses that are closely related tend to go into decline jointly. This could be taken to vindicate a lumping approach to semantic description, whereby whatever goes into joint decline makes up one major sense. That, however, would fail to capture the accompanying shifts in extensibility, as well as more subtle differences in the development of senses on a finer-grained analysis. For instance, the results from the case studies indicate that the periphery is lost before the core, starting from the sense furthest removed from the core.

These different points are illustrated by *awful*, where loss of the original core sense ('inspiring fear') coincides with the emergence of a new and highly frequent evaluative use ('unpleasant') that generates its own extensions (such as 'of poor quality'). As the old sanctioning sense loses its extensibility, senses derived from it (e.g. 'filled with fear') become obsolescent as well. The same pattern is seen in *about*, which has come close to losing its old sanctioning sense ('encircling') along with any immediately derived spatial senses (e.g. 'in the vicinity of'), at the same time as its semantic profile has come to be dominated by the cluster of mental connection senses organized around a new core sense 'concerning'. Interestingly, some senses – particularly the approximator use of *about* – appear relatively insulated from the effects of this reorganisation, possibly reflecting their different syntactic status. Finally, *so* offers the most complex picture. Its uses are so semantically and syntactically diverse that over the period investigated there is no point when its use could be said to be organized around a single sanctioning sense. Yet, semantic clustering in its uses is in evidence all the same during a major overhaul of its use in the twentieth century, when deictic and intensifying uses consistently decline, while causal uses are stable or increase.

These case studies do not solve the problems that beset RNT. They offer no ready answers where it comes to describing polysemy, deciding on the most likely sanctioning sense (or senses) in a radial network, or the exact nature of its motivational ties, or the number of separate senses it should include. Even so, they show that RNT makes plausible predictions about semantic change, in line with the assumptions that sense relations are motivational, that such motivation is mostly directional, and that it depends on the sanctioning status of a core sense. As such, these assumptions hold out the promise of a better understanding of the dynamics of semantic change.

Acknowledgements. We thank the reviewers for their constructive comments on an earlier version of the article. This research was supported by the Research Foundation Flanders (FWO), grant number 11PF524N. The first author is the lead author.

References

Allan, Kathryn. 2021. Metaphor, metonymy and polysemy: A historical perspective. In Augusto Soares da Silva (ed.), *Figurative language: Intersubjectivity and usage*, 287–306. Amsterdam: John Benjamins.

Baayen, Harald & Rochelle Lieber. 1991. Productivity and word-formation in English: A corpus-based study. *Linguistics* 29, 801–43.

Barðdal, Jóhanna. 2008. *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins.
Baumann, Andreas, Klaus Hofmann, Anna Marakasova, Julia Neidhardt & Tanja Wissik. 2023. Semantic micro-dynamics as a reflex of occurrence frequency: A semantic networks approach. *Cognitive Linguistics* 34(3–4), 533–68.
BNC Consortium, *The British National Corpus*, XML edition. 2007. Oxford Text Archive, http://hdl.handle.net/20.500.14106/2554 (accessed 15 May 2024).
Bolinger, Dwight. 1948. On defining the morpheme. *Word* 4, 18–23.
Brugman, Claudia. 1981. The story of 'over'. MA thesis, University of California, Berkeley.
Brugman, Claudia & George Lakoff. 1988. Cognitive topology and lexical networks. In Steven L. Small, Garrison W. Cottrell & Michael K. Tanenhaus (eds.), *Lexical ambiguity resolution*, 477–508. Palo Alto, CA: Morgan Kaufman.
Budts, Sara. 2020. On periphrastic *do* and the modal auxiliaries: A connectionist approach to language change. PhD dissertation, Antwerp University.
Bybee, Joan. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.
Cai, Yingying & Hendrik De Smet. 2024a. Are categories' cores more isomorphic than their peripheries? *Frontiers in Communication* 9, www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2024.1310234/full (accessed 9 September 2024).
Cai, Yingying & Hendrik De Smet. 2024b. In alternations, not all semantic motivation comes from semantic contrast. *Linguistic Vanguard* 10(1), 53–62.
De Smet, Hendrik. 2016. How gradual change progresses: The interaction between convention and innovation. *Language Variation and Change* 28(1), 83–102.
De Smet, Hendrik. 2017. Entrenchment effects in language change. In Hans-Jörg Schmid (ed.), *Entrenchment, memory and automaticity: The psychology of linguistic knowledge and language learning*, 75–99. Washington, DC: American Psychology Association.
De Smet, Hendrik, Susanne Flach, Jukka Tyrkkö & Hans-Jürgen Diller. 2015. *The Corpus of Late Modern English (CLMET), version 3.1: Improved tokenization and linguistic annotation.* KU Leuven, FU Berlin, U Tampere, RU Bochum. Available from perswww.kuleuven.be/~u0044428/clmet.htm (accessed 23 May 2024).
Evans, Vyvyan. 2004. *The structure of time: Language, meaning and temporal cognition*. Amsterdam: John Benjamins.
Evans, Vyvyan. 2005. The meaning of time: Polysemy, the lexicon and conceptual structure. *Journal of Linguistics* 41(1), 33–75.
Geeraerts, Dirk. 1997. *Diachronic prototype semantics: A contribution to historical lexicology*. Oxford and New York: Oxford University Press.
Geeraerts, Dirk. 2015. Sense individuation. In Nick Riemer (ed.), *The Routledge handbook of semantics*, 233–47. London: Routledge.
Gilquin, Gaëtanelle & Andrew McMichael. 2018. Through the prototypes of through: A corpus-based cognitive analysis. *Yearbook of the German Cognitive Linguistics Association* 6(1), 43–70.
Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.
Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
Gries, Stefan Th. 2010. Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. *The Mental Lexicon* 5(3), 323–46.
Györi, Gábor. 2002. Semantic change and cognition. *Cognitive Linguistics* 13(2), 123–66.
Hay, Jennifer. 2003. *Causes and consequences of word structure*. London: Routledge.
Hay, Jennifer & Harald Baayen. 2002. Parsing and productivity. In Gert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 2001*, 203–35. Dordrecht: Kluwer.
Heylen, Kris, Thomas Wielfaert, Dirk Speelman & Dirk Geeraerts. 2015. Monitoring polysemy: Word space models as a tool for large-scale lexical semantic analysis. *Lingua* 157, 153–72.
Hilpert, Martin & David Correia Saavedra. 2017. Using token-based semantic vector spaces for corpus linguistic analyses: From practical applications to tests of theoretical claims. *Corpus Linguistics and Linguistic Theory* 16(2), 393–424.
Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticalization. Opposite or orthogonal? In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components*, 21–42. Berlin: Mouton de Gruyter.
Huang, Danqing, Dirk Geeraerts & Weiwei Zhang. 2021. A diachronic analysis of the FIRE character. *Chinese Semiotic Studies* 17(1), 1–44.
Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.
Langacker, Ronald W. 1987. *Foundations of cognitive grammar*, vol. 1: *Theoretical prerequisites.* Stanford, CA: Stanford University Press.
Li, Jiangtian. 2024. Semantic minimalism and the continuous nature of polysemy. *Mind & Language* 39, 680–705.

Lopukhina, Anastasiya, Anna Laurinavichyute, Konstantin Lopukhin & Olga Dragoy. 2018. The mental representation of polysemy across word classes. *Frontiers in Psychology* 9, https://doi.org/10.3389/fpsyg.2018.00192 (accessed 2 September 2024).

Oxford English Dictionary (*OED*) Online. 2000–. 3rd edn (in progress). Oxford University Press. www.oed.com (accessed April 2024).

Petré, Peter. 2013. *EEBOCorp: Early English Books Online Corpus 1.0.* Available from https://lirias.kuleuven.be/handle/123456789/416330 (accessed 21 May 2024).

Petré, Peter. 2014. *Constructions and environments: Copular, passive, and related constructions in Old and Middle English*. Oxford: Oxford University Press.

Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin: Mouton de Gruyter.

Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch & Barbara B. Lloyd (eds.), *Cognition and categorization*, 27–48. Hillsdale, NJ: Lawrence Erlbaum.

Sandra, Dominiek & Sally Rice. 1995. Network analyses of prepositional meaning: Mirroring whose mind – the linguist's or the language user's? *Cognitive Linguistics* 6(1), 89–130.

Schmid, Hans-Jörg. 2018. Unifying entrenched tokens and schematized types as routinized commonalities of linguistic experience. *Yearbook of the German Cognitive Linguistics Association* 6, 167–82.

Taylor, John R. 1995. *Linguistic categorization: Prototypes in linguistic theory*, 2nd edn. Oxford: Clarendon Press.

Tyler, Andrea & Vyvyan Evans. 2001. Reconsidering prepositional polysemy networks: The case of *over*. *Language* 77(4), 724–65.

Tyler, Andrea & Vyvyan Evans. 2003. *The semantics of English prepositions: Spatial scenes, embodied meaning, and cognition*. Cambridge: Cambridge University Press.

Cite this article: Ceuppens, Hilke and De Smet, Hendrik. 2025. Tied together, fried together: understanding semantic extensibility through semantic loss. *English Language and Linguistics* 1–17. https://doi.org/10.1017/S1360674325000206