María José López-Couso*

Continuing the dialogue between corpus linguistics and grammaticalization theory: Three case studies

DOI 10.1515/cllt-2015-0069

Abstract: The article shows that significant benefits can be gained from the integration of corpus linguistics and grammaticalization theory, two subfields which, despite sharing considerable common ground, tended to remain as separate areas of linguistic analysis until quite recently. Making use of diachronic and contemporary corpora of English, such as the Helsinki Corpus, ARCHER, and COCA, the article illustrates how standard corpus practices can indeed contribute to our understanding of grammaticalization and related processes of language change. The selected case studies deal with the origin of existential there, the development of like-parentheticals in contemporary American English, and the history of the marker of expository apposition namely.

Keywords: corpus linguistics, grammaticalization, (inter)subjectification, existential there, deictic there, like-parentheticals, comparative complementizer like, marker of expository apposition namely, particularizer namely, bridging context

1 Introduction

The last couple of decades have witnessed an increasing awareness of the benefits to be derived from the integration of corpus linguistics and grammaticalization theory, two areas which until quite recently had tended to remain separate. Corpus linguistics provides sound empirical methodology for the recognition and documentation of grammaticalization processes, by making use of computerized corpora and relying on established statistical practices, especially those developed during the latter part of the twentieth century. In turn, grammaticalization theory helps to bring corpus linguistics beyond the purely statistical domain, “liberating” it from the stigma of being seen as nothing more than “a cemetery of numbers, – an incoherent compilation of uninterpreted and hence pointless statistics” (Mair 2004: 139).

*Corresponding author: María José López-Couso, Department of English and German, University of Santiago de Compostela, E-15782 Santiago de Compostela, Spain, E-mail: mjlopez.couso@usc.es
The present article aims to show how corpus linguistics can effectively be related to the concerns of grammaticalization theory, thus contributing to our understanding of processes of language change like grammaticalization, subjectification, and intersubjectification. The article is organized as follows. Section 2 provides some background on the relation between corpus linguistics and grammaticalization studies over the past two decades. Making use of historical and present-day computerized corpora, such as the *Helsinki Corpus of English Texts* (HC), *A Representative Corpus of Historical English Registers* (ARCHER), and the *Corpus of Contemporary American English* (COCA, Davies 2008–), Section 3 moves on to illustrate the interplay of these two linguistic subfields by means of three case studies from my recent and ongoing research on linguistic change in different domains: existential *there* (3.1), *like*-parentheticals (3.2), and the marker of expository apposition *namely* (3.3). Finally, Section 4 offers some concluding remarks.

## 2 On the relation between corpus linguistics and grammaticalization theory

Although it has been a century ago now since Antoine Meillet first recognized the central importance of grammaticalization processes in the domain of language change, coining the term “grammaticalization” and indeed devoting an entire work to the topic (Meillet 1912), grammaticalization has become a major area in linguistic studies only during the last three decades or so, when it has established itself as “an important (but also controversial) concept in general and typological linguistics and a prominent type of explanation in historical linguistics” (Lindquist and Mair 2004: ix). The methodology used in the most influential works on grammaticalization from the 1990s, including Hopper and Traugott’s (1993) textbook and the collections of papers edited by Heine et al. (1991) and Traugott and Heine (1991), is largely of a qualitative, rather than a quantitative, nature. Moreover, many of the grammaticalization theorists of the time who make use of corpus data for their analyses typically resort to self-compiled and relatively small corpora rather than to large computerized databases (see Mair 2004: 122). Consider, for example, the way in which Lehmann describes the data used for his 1991 study on “Grammaticalization and related changes in Contemporary German”:

> For the past three years or so, I have been gathering data on the present topic. They are instances of ongoing changes or current fashions, *occasionally encountered and unsystematically noted down*. Some of them are tied up with my own dialect environment, which is
Northwest Germany. For several of the phenomena to be mentioned below, I have no spontaneous data at all. (Lehmann 1991: 494–495; emphasis mine)

It must be noted, however, that even though the methodology used in most of these classic works on grammaticalization is “largely qualitative rather than primarily quantitative and corpus-based” (Nevalainen 2004: 4), on occasion they can come close to corpus-based approaches to language change. Thus, for example, Hopper and Traugott (1993) stress the importance of studying “patterns of usage, as reflected by the frequency with which tokens of these structures may occur across time” (1993: 59; 2003: 67). More significantly, these authors close their discussion of relative degrees of grammaticalization of vector verbs in Hindi-Urdu and Kashmiri in the following way:

*Statistical evidence is a valuable tool in providing empirical evidence for unidirectionality.*

For diachronic studies access to texts of comparable genres over a fairly long period is needed. It is only in a few languages that we are fortunate enough to have this kind of textual history. And it is for only a small subset of these languages that we have any statistical studies at all of the development of grammatical items. There is an urgent need for additional reliable statistic studies of a variety of phenomena in which early grammaticalization appears to be involved... More work is necessary to diagnose grammaticalization in its early stages and to develop the kinds of statistical parameters which will reveal it. (Hopper and Traugott 1993: 111–112; emphasis mine)

Interestingly, this paragraph was omitted from the 2003 edition of Hopper and Traugott’s monograph, which seems to imply that the pressing need for statistical studies of grammaticalization phenomena mentioned by the authors back in 1993 had already been satisfied (at least partially) ten years later.

It is no wonder, then, that Hopper and Traugott (1993) appear to echo corpus linguists on these and other occasions. In an excellent article discussing the relation between corpus linguistics and grammaticalization theory, Mair (2004) shows that there is significant common ground shared by these two linguistic subfields:

- Both approaches give priority to the study of utterances in their discourse contexts rather than abstract systems of underlying rules.
- Both emphasize the importance of frequency data and statistics.
- Both agree that transitions between grammatical categories are gradient rather than abrupt, and that grammatical form and meaning are interdependent rather than constituting separate and autonomous domains.
- Both, finally, became “hot” in linguistics again in the late nineteen seventies and early nineteen eighties after decades of relative neglect. (Mair 2004: 121)

In spite of sharing a considerable amount of common ground, until quite recently corpus linguistics and grammaticalization theory tended to remain as
separate areas of linguistic study. As a revealing example of the early divide between the two subfields, Mair (2004: 122) cites Olofsson’s (1990) article on the development of the preposition *following*, meaning ‘after, as a result of a particular event’, as in (1):

(1)  *He took charge of the family business following his father’s death.*

On the basis of data from the Brown and LOB corpora, Olofsson discusses the grammaticalization of the participle *following* from a verbal form to a preposition, though the term “grammaticalization” is never mentioned in the article. Twenty-one years later, Olofsson published a follow-up paper, “Prepositional *following* revisited” (2011). Here, the term “grammaticalization” is used about a dozen times, twice on the first page of the introduction.

Olofsson’s (1990) article is not unusual as an illustration of the huge gap that existed until quite recently between corpus linguistics and grammaticalization theory. However, a considerable number of studies published over the last twenty years testify to the increasingly closer collaboration between the two subfields. Major landmarks in this productive dialogue include the following: Mair’s (1994) article “Is see becoming a conjunction? The study of grammaticalisation as a meeting ground for corpus linguistics and grammaticalisation theory”; the collection of papers *Grammaticalization at work: Studies of long-term developments in English* (Rissanen et al. 1997) on grammaticalization processes in earlier stages of the English language; the symposium “Corpus research on grammaticalization in English” organized by Hans Lindquist at Växjö University (Sweden) in 2001; and the volume *Corpus approaches to grammaticalization in English*, edited by Lindquist and Mair (2004), which contains most of the papers presented at the symposium, together with several additional articles at the invitation of the editors. Among the most notable recent examples of the ever-increasing convergence of grammaticalization studies and corpus linguistics are Hilpert’s (2008) analysis of Germanic future constructions and several of the contributions in Traugott and Trousdale’s volume *Gradience, gradualness and grammaticalization* (2010), in particular those by De Smet (on the *for* ... *to*-infinitive and the phrasal verb particles *out* and *forth*), Denison (on category changes), and Patten (on *it*-cleft constructions).

---

1 The main findings of Olofsson’s (2011) study are that the prepositional use of *following* was established earlier than had previously been noted and that its frequency in British English has tripled over the last fifty years.
3 Illustrating the dialogue between corpus linguistics and grammaticalization studies

The present section offers further proof of this fruitful dialogue between corpus linguistics and grammaticalization studies by means of the analysis of three grammaticalization processes from both earlier and Present-day English.

3.1 Looking for parallels between ontogenetic and diachronic grammaticalization – Case study 1: Existential there

The first case study here concerns the grammaticalization of existential there. It is generally agreed that existential there developed from the distal deictic adverb there, meaning ‘in/at that place’ (see, among many others, Lyons 1967: 390; Bolinger 1977: 91–92; Breivik 1983, 1997: 32; Johansson 1997; Pfenninger 2009: 49–53), through processes of grammaticalization and reanalysis, probably before the first written records of English (see Breivik 1977: 346). If the beginning of these processes of change indeed pre-date Old English times, lack of data for pre-literary periods may make it difficult (if not impossible) to plot the split of existential there from locative there. Nevertheless, the Old English period clearly represents a transitional stage in the historical development of existential there. Notice, for instance, that existentials with no introductory element represented the most common existential pattern at this early stage (see Breivik 1983: 278, 319; López-Couso 2006: 182, among others), while there-existentials became the default type of existential construction only during the Middle English period (see Breivik 1983: 320–321).

Looking at the ontogenetic development of existential there, and using longitudinal corpus evidence from the CHILDES archive (MacWhinney 1995), Christopher Johnson (1999, 2001, 2005) has shown that, in the learning process, children use locative there as a source for existential there, via overlap utterances which share formal and semantic-pragmatic properties of either construction. Taking Johnson’s analysis as a starting point, I examine in López-Couso (2011) whether the developmental relation between deictic there-constructions

\[2\] For an alternative non-locative account of the origin of existential there, see Davidse (1992).
and there-existentials explained by Johnson’s theory of “constructional grounding” or “developmental reinterpretation” for child language acquisition is valid also in diachronic terms.

My analysis of the prose texts of the Old and Early Middle English sections of the HC\(^3\) suggests that the emergence of existential *there* in the history of the language ran parallel to the three developmental stages proposed by Johnson for Child English (see Johnson 1999: 95–114; 2001: 131–134).

**Stage 1.** At the initial stage, both ontogenetically and diachronically, *there* is a distal deictic adverb, meaning ‘in/at that place’, as shown in (2a–b):

(2) a. *There’s Mommy.* (CHILDES; Naomi, age: 1;10)
   
b. *ðar* was se cing gehaten Sæbyrht. Ricolan sunu.
   
   there was the king called S. R.’s son
   
   ‘There was the king named Sæbyrht, the son of Ricola.’
   
   (HC; O2, Two of the Saxon Chronicles Parallel, R 604.3)

**Stage 2.** Over the course of time, locative *there* begins to be found in overlap contexts, which allow both a deictic and an existential reading, as in (3a–b). Johnson describes such contexts as “overlap deictics”, i.e., statements in which *there* performs an existence-informing function in addition to its original pointing-out function. Historically, as mentioned above, the transition from Stage 1 to Stage 2 is most likely to have taken place before the Old English period.

(3) a. *There’s cup for Mom.* (CHILDES; Naomi, age: 2;5)
   
b. Ac þa strengstan weras wuniaþ on þam lande & micele burga
   but the strongest men live in the land and great cities
   
   *ðær* synd & mærlice geweallode:
   there are and splendidly walled
   
   ‘But the strongest people live in the land and there are large and splendidly fortified cities there:
   
   (HC; O3, The Old Testament, Numbers 13.29)

A particularly relevant subtype of overlap deictics identified by Johnson in Child English is the so-called “double-locative overlap deictic” construction (Johnson 1999: 85), illustrated in (4a) below; it is characterized by the co-occurrence of

3 Periodization of the relevant sections of the HC is as follows: O1 (–850); O2 (850–950); O3 (950–1050); O4 (1050–1150); M1 (1150–1250); M2 (1250–1350).
there with an additional locative expression in the clause. Constructions of this kind seem to play a prominent role in the process of change in first language acquisition, since children need to look for a new (non-locative) function of there, which would otherwise be perceived as redundant. As in early child speech, double-locative overlap constructions seem to have served as bridging contexts (Heine 2002) in the diachronic grammaticalization of existential there (cf. (4b)):

(4)  
   a. *There’s a table on the house.* (CHILDES; Nina, age: 2;1)  
   b. *for þær wæs an forehus æt þære cyrcan duru*  
      for there was a porch at the church door  
      ‘for there was a porch at the church door’  
      (HC; O4, An Old English Vision of Leofric, Earl of Mercia, 31)

Stage 3. After the overlap stage, children produce the first unambiguous instances of there-existentials. Notice, for example, that in (5a) a locative reading of there would not make sense, given the co-occurrence of the proximal deictic here in the same clause. Historically, the differentiation between locative there and existential there is finally accomplished, most probably in post-Old English times (cf. the Early Middle English example in (5b)). Interestingly, it seems that unequivocal instances of existential there appear earlier in affirmative clauses both in Child English and in the history of the language, and later spread to negated contexts (cf. (5c–d)):

(5)  
   a. *There’s money in here.* (CHILDES; Peter, age: 2;5)  
   b. *ʒeʃ pu get wite wult hwucche wihtes þear beon þear as*  
      if you yet know want which creatures there are where  
      *al þis blisse is … Ich þe ontswerie;*  
      all this bliss is I you answer  
      ‘If you want to know what kind of creatures there are where all this  
        happiness is … I answer you;’  
      (HC; M1, The Katherine Group, 40)  
   c. *There’s no fire.* (CHILDES; Peter, age: 2;7)  
   d. *And þou forseȝest alle myn waies, for þer nis no worde*  
      and you know all my ways for there not.is no word  
      *in my tunge.*  
      in my tongue  
      ‘And you know all my ways, for there is no word in my tongue.’  
      (HC; M2, The Earliest Complete English Prose Psalter, Psalm 138.3)
Figure 1 summarizes the development of the overlap deictics (cf. (3)), double-locative overlap deictics (cf. (4)), and unambiguous existentials (cf. (5)) in the Old and Early Middle English sections of the HC:

![Figure 1: Overlap deictics, double-locative overlap deictics, and unambiguous existentials in the prose texts of the Old and Early Middle English sections of the HC (normalized frequencies per 100,000 words).](image)

The close similarity between the ontogenetic and diachronic processes of grammaticalization of existential *there* just outlined is probably more than mere coincidence. Similar corpus-based studies may reveal the existence of further parallels in the way grammaticalization works in child language acquisition and in the history of the language.

### 3.2 Recognizing and documenting incipient or ongoing grammaticalization – Case study 2: *Like*-parentheticals

My second case study brings us forward in time to the present day. From as far back as the early 1990s, one of the controversial issues in grammaticalization has been whether it is possible to recognize and document incipient and ongoing grammaticalization processes. The question was addressed by Christian Lehmann (1991), who reviewed, from the perspective of grammaticalization, a number of constructions in both the nominal and verbal spheres in contemporary non-literary German; these constructions were in vogue at the

---

4 In view of the potential influence of syllable count and alliteration on the presence/absence of existential *there* in verse works in Early English, only prose texts have been considered for the analysis.
time, and although not accepted by standard grammars they did show some of the features characteristic of grammaticalization. In the closing section of his article, Lehmann admits that the study of ongoing grammaticalization processes “is subject to a serious problem of verification” and that “it is next to impossible to know which of the changes that speech habits currently exhibit are synchronic manifestations of ongoing genuine language change, and which of them are but ephemeral fashions” (Lehmann 1991: 532).

In spite of this note of skepticism – and similar ones – the last few years have seen a number of studies on various grammaticalization phenomena which have convincingly demonstrated that, at least for English, the use of corpora can help to pinpoint incipient or ongoing grammaticalization. Evidence here can be found in the analysis of the emergence of the epistemic/evidential parentheticals looks like, seems like, and sounds like in contemporary American English, which I investigate in collaboration with my colleague Belén Méndez-Naya (López-Couso and Méndez-Naya 2014a, 2015). Examples of the pattern we are interested in are given here:

(6) a. We now, it seems like, get an annual letter. (COCA; 2005, SPOK, CNN_Dobbs)

b. “You looked good out there.” “Thanks.” “Made a couple of friends, looks like.” (COCA; 2003, FIC, Highlights)

c. He was mentally torturing you, it sounds like. (COCA; 2008, SPOK, CBS 48Hours)

Notice that this parenthetical type differs from prototypical clausal parentheticals, such as I think, I believe, you know, and it seems (López-Couso and Méndez-Naya 2014b, 2014c), in that the conjunction like has become bonded to the verb as a result of a process of fusion, one of the characteristics Brinton and Traugott (2005: 27–28) associate with grammaticalization processes (see also Brinton 2008: 50).

For the analysis of the parentheticals under discussion we make use of the Corpus of Contemporary American English (COCA, Davies 2008–), which contains

---

5 Among other topics, Lehmann discusses the development of complex prepositions like im Zuge ‘by, during, in’ and im Wege ‘by (way/means of)’, modal and function verbs, intensification (mehr als, zunehmen), verb-second position in sub-clauses, and verb-first in main clauses.

6 As Lindquist and Mair (2004: x) put it, “There is now a corpus-linguistic working environment for English which comprises corpora which, taken together, are a body of evidence which is not only massive in size but also of extremely high quality because of its differentiated coverage of written and spoken material, of past and present stages of the language, and of different text types and genres. It is thus expected that grammaticalization processes can be studied in much more detail than would be possible for most other languages, and that the results thus obtained will lead to a refinement of the theoretical model”.

about 450 million words of material from 1990 to the present day, and is therefore an excellent source for the study of recent and ongoing developments in American English. Figure 2 shows the distribution of relevant instances in the corpus from the early 1990s to the year 2009, arranged in five-year periods. The data show a clear increase in the frequency of like-parenthetical structures with look and sound over the last two decades. By contrast, like-combinations with seem remain rather stable over time, probably because of the high frequency of the bare pattern it seems, which constitutes the default parenthetical type with this predicate (see López-Couso and Méndez-Naya 2014a, 2014c).

In order to determine whether the development of the parentheticals looks like, seems like, and sounds like can be conceptualized in terms of processes such as grammaticalization, subjectification, and/or intersubjectification, we have compared them with related complement constructions featuring the predicates look, seem, and sound and the “comparative complementizer” like of the type shown here:

(7) a. **It seems like** he doesn’t want any vetting going on. (COCA; 2008, SPOK, Fox Hannity and Colmes)
   b. **It looks like** we’re going to spend the night where we are. (COCA; 2008, MAG, Field and Stream)
   c. **It sounds like** we need a review of the hospital’s notification policy. (COCA; 2009, NEWS, The Denver Post)

7 In addition, COCA is evenly divided into five different genres: spoken (mainly transcribed conversation from TV and radio), fiction, popular magazines, and academic journals from different fields, which allows for the study of a given phenomenon across text-types. For information on the corpus, see http://corpus.byu.edu/coca/.
8 On so-called comparative complementizers (as if, as though, and like), see López-Couso and Méndez-Naya (2012a, 2012b).
Looking closely at the morphosyntactic and semantic-pragmatic characteristics of the *like*-parentheticals attested in COCA, it can be seen that these combinations illustrate a number of features which are typical of grammaticalization and related processes of language change. For example, they have undergone a process of decategorialization: originally complement-taking clauses (see (7) above), they are downgraded to parenthetical constructions which no longer show a complete syntactic structure (see (6a–c)). Once they have achieved parenthetical status, these combinations become syntactically and prosodically independent and may therefore occupy different positions in the clause (e.g. medial (6a) and final (6b–c)), thus coming close to adverbs. The coexistence in contemporary American English of the different related constructions (the complement-taking clause, the parenthetical structure, and the quasi-adverb) points at another of the criterial features of grammaticalization, namely layering (see Hopper 1991).

Hand in hand with decategorialization, the parentheticals at issue show a high degree of morphosyntactic fixation; this is in clear contrast to the related complementation patterns, which display a wider range of variability. For one, as regards tense, aspect, and mood (TAM) distinctions, the *like*-parentheticals in our data show a marked tendency to occur in the present tense (75.6%; see the examples in (6) above) and, to a much lesser extent, in the preterite (24.4%), as shown in (8). On the other hand, when used as matrix predicates in complementation structures, the verbs *look*, *seem*, and *sound* not only occur in the present and the preterite, but also in combination with modals and other auxiliaries (cf. (9a–b)) as well as in non-finite forms in catenative constructions (see (9c)):

(8) “Pamela?” Wim’s voice had thinned out, *seemed like*. “Sweetheart, we’re here,” he said. (COCA; 1999, FIC, Dodd *The mourner’s bench: A novel*)

(9) a. *It may seem like* he has vanished from Chicago life, like a rat in witness protection, (COCA; 2007, NEWS, Chicago)

b. *In fact, we were the first ones to report earlier today that according to a White House source, it did look like* those talks broke down. (COCA; 1998, SPOK, Fox_HC)

c. “*You make it sound like* we’re living in a fairy tale,” Mack said. (COCA; 2001, FIC, Bk: BeachClub)

---

9 On these and other features characterizing parentheticals, see Dehé and Kavalova (2007), Kaltenböck (2007), and Brinton (2008).

10 Additional examples of the development of original matrix clauses in complement constructions into epistemic adverbs are *methinks* in earlier English (see López-Couso 1996; Palander-Collin 1999; Wischer 2000) and Danish *måske* and Danish and Swedish *kanske* ‘maybe’ (see Boye and Harder 2007: 591), among others.
Besides their limited range of TAM distinctions, like-parentheticals also show a significant restriction as regards polarity. Our data from COCA indicate that these parenthetical constructions tend to occur in the affirmative (99% of all relevant instances), while negated like-parentheticals are only sporadically attested; sentence (10) below is one of these exceptional instances. In all such cases, the anchor clause is also negative, in accordance with the claim that a negative parenthetical can only occur in combination with a negative host (Kaltenböck 2007: 37; Schneider 2007: 149–150).

(10) A hospital spokesman says they’re staying there for a couple more days, so he won’t be lighting any cauldron, it doesn’t seem like. (COCA; 2004, SPOK, NPR_Morning)

Further indicators of the high degree of morphosyntactic fixation of the like-parentheticals are their inability to show adverbial modification or to occur in interrogative structures. As shown in (11a–b), these two patterns are available for the related complementation structures:

(11) a. It really looked like he was not going to make it. (COCA; 2002, SPOK, CBS SixtyII)

b. “Why does it seem like I’m the only one who got punished?” (COCA; 2006, FIC, BkSF:DialLLoser)

Finally, even though the predicate and the conjunction making up the like-parentheticals remain orthographically separate, they do show fusion (see above): the predicate and like are bonded, and come to be understood as constituting a single chunk, thus disallowing intervening material. By contrast, the occurrence of intervening elements between the predicate and like is possible in the related complementation structures, as shown in (12):

(12) and so many people to look at that it seemed almost like they weren’t narrowing it down. (COCA; 2009, SPOK, CNN_Cooper)

In addition to the morphosyntactic features discussed so far, like-parentheticals also undergo interesting semantic-pragmatic changes. On the one hand, when used in the combinations currently under discussion, the predicates look, seem, and sound show reduced semantic content, their sensory meaning being bleached. On the other hand, lacking propositional semantics, like-parentheticals specialize in subjective and intersubjective meanings and functions (see López-Couso 2010). Thus, they can be used, for instance, to mitigate the
speaker’s commitment to the truth of a proposition. Consider in this connection our earlier examples (6a–b), where the parentheticals *it seems like* and *looks like* indicate that the speakers do not fully commit themselves to what is said in the respective anchor clauses, probably suggesting a lack of definite evidence for their statements. Moreover, in conversation (see (13) below), *like*-parentheticals may also serve intersubjective functions as face-saving devices: they soften the speaker’s assertion and hence allow the interlocutor(s) to disagree:  

(13) *Are you happy with the way it’s turning out? Bush ran out of New Hampshire the other night for the debate. He’s scared to debate, it looks like.* (COCA; 1999, SPOK, Fox_Drudge)

### 3.3 Corpora as a source for qualitative analysis of grammaticalization phenomena – Case study 3: Optional marker of expository apposition *namely*

The development of the *like*-parentheticals discussed in Section 3.2 shows that large databases like COCA can help us to identify, describe, and analyze incipient and ongoing grammaticalization. However, not all grammaticalization processes “leave a statistical imprint” (Mair 2004: 133) in corpora in the way existential *there* or *like*-parentheticals do. This is especially problematic for historical analyses of low-frequency phenomena such as the one in my third case study:  

12 the development of the optional marker of expository apposition *namely*, illustrated in (14).

(14) *How can a solution be found to the current disease of contemporary society, namely the international economic crisis?* (from Quirk et al. 1985: 1307)

According to the *Oxford English Dictionary* (OED) and the *Middle English Dictionary* (MED), the origin of the adverb *namely* is found in the combination of the noun *name* and the derivational suffix -ly (see OED s.v. *namely*; MED s.v. *nam(e)li* adv.). A look at the entries for *namely* in these two dictionaries reveals the existence of three different meanings for the form over the course of time. The original meaning of *namely* seems to have been that of ‘particularly,

---

12 For a particularly interesting corpus-based study of the grammaticalization of low-frequency phenomena, see Hoffmann (2004) on the development of low-frequency complex prepositions of the type *by dint of*, *in conformity with*, and *in search for.*
especially, above all’ (see OED s.v. *namely* adv. 1; MED s.v. *nam(e)li* adv. 1). The first occurrence of the form I have been able to trace (example (15) below) dates from the late twelfth-century text of the *Lambeth Homilies*:

(15) *Sunnedei ah efri cristenne Mon noneliche to chirche cume.*

Sunday should every Christian man especially to church come

‘On Sunday every Christian should especially come to church.’

c1175, *Lambeth Homilies* 139; OED s.v. *namely* adv. 1a; MED s.v. *nam(e)li* adv. 1a).

Similar instances survived into the Early Modern English period, but became obsolete after 1700. Example (16) is the last occurrence of this kind recorded in the OED:

(16) *Returning thanks ... for many blessings and favors ... And, namely, for the enjoyment of the Gospel.* (1700, *New Hampsh. Prov. Papers* (1868) III. 327; OED s.v. *namely* adv. 1a).

In examples like (15) and (16), *namely* is used as a particularizer, whose function is to “restrict the application of the utterance *predominantly* to the part focused” (Quirk et al. 1985: 604; original emphasis).

A second possible meaning of *namely* in earlier stages of the language is ‘at least, at any rate’ (see OED s.v. *namely* adv. 2; MED s.v. *nam(e)li* adv. 3), as in example (17) from the MED. It seems, however, that this use of *namely* was rather rare and did not survive the Middle English period.

(17) *Assenteþ to me, nameliche, in þre þinges, þif ðe willeþ nouȝt assente to me and þe oþere.*

‘Agree with me at least on three things, if you do not want to agree with me on the other.’

(a1387, Trev. *Higd.* 5.407; MED s.v. *nam(e)li* adv. 3a)

During Late Middle English, *namely* acquired a third possible meaning: ‘to wit, that is to say, videlicet’ (see OED s.v. *namely* adv. 3; MED s.v. *nam(e)li* adv. 2). Example (18), dating from the mid-fifteenth century, is one of the earliest attested unambiguous occurrences of *namely* with this meaning:

(18) *Beryn cam nat þere, Namelich in-to the place there his modir lay.*

‘Beryn did not come there, that is to say to the place where his mother was lying.’

c1460, *Beryn* 1049; MED s.v. *nam(e)li* adv. 2a).
In such cases, *namely* is not a particularizer but rather functions as an optional marker of apposition (see Quirk et al. 1985: 1307–1316; Meyer 1992: 97), which can be omitted without affecting the grammaticality of the sequence. More specifically, in (18) *namely* marks a relation of expository apposition, where “the units are referentially or semantically equivalent” (Meyer 1992: 97). This is the only meaning and function that the adverb *namely* retains in Present-day English.

Figure 3 shows the distribution of *namely* in the HC and ARCHER 3.1 according to function. As can be seen, in the Middle English material *namely* is recorded only as a particularizer, and in this function it is still sporadically attested by the end of the Early Modern English period. On the other hand, *namely* as a marker of expository apposition occurs in the corpus data only from the early seventeenth century onwards. This latter function is the only one attested in the Late Modern and Present-day English data from ARCHER.

![Figure 3: Functions of *namely* in the HC (M1–E3) and ARCHER 3.1 (eighteenth- to twentieth-century British English) (normalized frequencies per 100,000 words).](image-url)

Examples (19) and (20) from the HC show *namely* in its function as a particularizer and as a marker of expository apposition, respectively:

(19) *for no doute he shal fynde ful manye biblis in Latyn ful false, if he loke manie, nameli newe;*
   ‘for no doubt he shall find many Bibles in Latin that are completely false, if he examines many, especially the recent ones;’
   (HC; c1388, John Purvey, *The Prologue to the Bible*, I, 58)

---

13 Periodization of the relevant sections of the HC is as follows: M1 (1150–1250); M2 (1250–1350); M3 (1350–1420); M4 (1420–1500); E1 (1500–1570); E2 (1570–1640); E3 (1640–1710). The first ARCHER subperiod, 1650–1699, has been disregarded here due to its (partial) overlap with the last HC subperiod (1640–1710). Five instances of *namely*, all of them as an optional marker of expository apposition, have been recorded in this time span in ARCHER.
(20) ... let me assure you, that even with the other part that is wont to flye away, 
(namely the Flowers) and Antimonial Glass may without an addition of 
other Ingredients be made. (HC; 1675–6, Robert Boyle, Electricity & 
Magnetism, 23)

_Namely_ is also attested sporadically in the Early Modern English data from the 
HC in the combination _as namely_, which functions as an optional marker of 
exemplifying apposition meaning ‘for example’ (cf. OED s.v. _namely_ adv. 3b), as 
shown here:¹⁴

(21) ['\[1. Fri.\]'] when some wold willingly have goen to Plymmowth, some, _as namely_ M. Whood, desyred that at least the Frances myght turne in thither 
and fet M. Hawkins, 

(HC; 1582, Richard Madox, _An Elizabethan in 1582: The Diary of Richard 
Madox, Fellow of all Souls_, 140)

Two important features distinguish the particularizer _namely_ from the marker of 
expository apposition _namely_:

(i) As happens with other items with a particularizing function, the particu-
larizer _namely_ can occupy different positions in relation to the part of the 
utterance that is focused. Although it most commonly occurs in pre-position, 
as shown in (19) above, it can also occur postposed, as in (22) below. 
Postposition, however, seems a highly marked option; in my data it is only 
attested in four instances out of a total of 55 (7.3%).

(22) _His syns sal þan be shewed ful many, Als I tald byfor in þe thred part namly_. 
‘His sins shall then be shown to be many, as I told before in the third part 
especially.’ 

(HC; M3, _The Pricke of Conscience_, 80)

(ii) When occurring in pre-position, the particularizer _namely_ can be preceded by 
the conjunction _and_. An analysis of the HC reveals that _and_, though optional, 
very frequently occurs with the particularizer _namely_ in both Middle and Early 
Modern English (35 out of 51 relevant cases; 68.6%). Our earlier instance (16) 
and example (23) below from the HC illustrate this pattern:

---

¹⁴ In contrast to expository apposition, in exemplifying apposition the units are referentially 
and semantically only semi-equivalent (see Meyer 1992: 97). Markers of exemplifying apposition 
include _for example, for instance, say, and e.g.,_ among others (see Quirk et al. 1985: 1307; Meyer 
Through univerbation, the combination and namely sometimes resulted in the form anameli/anamely, as in (24) from the MED. No instances of this kind, however, have been attested in my corpus data.

Every good knyght shulde loue and prayse euer vertuous persone, anamely a woman in strong vertue of wytte and concyens.

‘Every good knight should always love and praise a virtuous person, especially a woman who is strong in virtue of thought and consciousness.’

(c1450, Scrope Othea 27; MED s.v. nam(e)li adv. 1a)

The core syntactic, semantic, and pragmatic features of namely in its two basic functions over time are summarized in Table 1.

Table 1: A comparison of the two basic uses of namely through history.

<table>
<thead>
<tr>
<th>Particularizer</th>
<th>Marker of expository apposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantics</td>
<td>‘particularly, especially’</td>
</tr>
<tr>
<td>Obligatory presence</td>
<td>Yes</td>
</tr>
<tr>
<td>Position</td>
<td>Preposed and postposed</td>
</tr>
<tr>
<td>Preceded by and</td>
<td>Yes (optionally)</td>
</tr>
</tbody>
</table>

Given that optional markers of apposition are rather infrequent in discourse, tracing the origin of the appositive marker namely through statistical analysis of historical data can turn out to be a difficult task. However, “corpora can be (and should be) sources for qualitative as well as quantitative analyses” (Lindquist and Mair 2004: xiii).

A close reading of the very few attestations of namely in the HC suggests a possible grammaticalization route from the particularizer namely to the optional appositive marker namely. A comparison of the two uses of the form as shown in Table 1 suggests that this change involved both semantic and functional

15 For example, they represent just 3% of the total of appositive constructions in Meyer’s analysis of apposition in contemporary English (1992: 97–98).
discontinuity. As happens with other particularizers, \textit{namely} in its particularizing function could occupy different positions in relation to the part of the utterance that is focused. Although it most commonly occurred in pre-position (see (19) above), it could occasionally appear postposed, as in (22). As shown below, instances of the latter type may have served as bridging contexts (Heine 2002) or critical contexts (Diewald 2002)\textsuperscript{16} in the development of \textit{namely} from particularizer to optional appositive marker. Consider in this vein example (25), from Dan Michel’s early fourteenth-century text \textit{Ayenbite of Inwyt}, a translation from the late thirteenth-century French work \textit{La Somme le Roi}, a manual of moral instruction written by Friar Lorens d’Orleans at the request of King Philip the Bold. Here, \textit{namely} functions as a particularizer, it is not preceded by \textit{and}, and is postposed to the part of the utterance which is focused (\textit{pri þing}). Example (25) can be compared with (26), which offers the same passage in \textit{The Book of Vices and Virtues},\textsuperscript{17} a later independent translation of the same French original. The punctuation of this version suggests that the change from particularizer to marker of expository apposition could already have taken place by the mid-fifteenth century, possibly as a result of faulty scribal analysis and resegmentation, as suggested by Pahta and Nevanlinna (2001). Example (26) is the only instance in the Middle English section of the HC where \textit{namely} would allow an interpretation as an optional marker of expository apposition.

(25) \textit{Dis word oure. ous têkþ to hatye \textit{pri þing} nameliche. Prede. wrepþe. and auarice.} \\
‘This word “ore” teaches us to hate three things in particular. Pride, hate, and avarice.’ \\
(HC; 1340, \textit{Ayenbite of Inwyt} I, 102).

(26) \textit{Dis word ‘oure’ têcep vs to hate \textit{pre þinges, namely: pride, hate, couetise.}} \\
‘This word “ore” teaches us to hate three things, namely: pride, hate, and avarice.’

(HC; c1450, \textit{Vices and Virtues} 4, 98).

It appears, then, that the development of \textit{namely} as seen here – from a marker of particularization in Early Middle English to an optional marker of expository apposition in Late Middle English – shows that corpora may “provide insights into grammaticalisation processes that go beyond the statistical” (Mair 2004: 138).

\textsuperscript{16} Diewald (2002: 103) defines critical context as that which is “characterized by multiple structural and semantic ambiguities and thus invites several alternative interpretations, among them the new grammatical meaning”.

\textsuperscript{17} The edition used for \textit{The Book of Vices and Virtues} in the HC is that by W. N. Francis (1942) for the Early English Text Society (EETS, number 217).
4 Closing remarks

The foregoing discussion has shown that significant benefits can be gained from the integration of corpus linguistics and grammaticalization theory. The three case studies described above as a means of illustrating the productive dialogue between these two linguistic subfields confirm that the use of computerized corpora and of standard corpus practices may prove relevant for our understanding of grammaticalization and related processes of language change. This is particularly so for specific problematic areas in grammaticalization studies. Thus, the comparison between the ontogenetic and the historical processes of grammaticalization of existential there (see Section 3.1), on the basis of data from the CHILDES archive and the HC, suggests that it is indeed possible to approach the study of grammaticalization phenomena which seem to have started before the earliest written records of English. Similarly, the analysis of like-parentheticals provided in Section 3.2 demonstrates that, even without the benefit of hindsight, ongoing grammaticalization can be recognized and documented with the help of large databases of contemporary English like COCA. Finally, the examination of the historical development of the marker of expository apposition namely (see Section 3.3) shows that, in addition to the unquestionable relevance of any statistical findings they may yield, corpora are also valuable sources of data for the qualitative analysis of low-frequency items and constructions.

Acknowledgement: I would like to thank an anonymous reviewer, the editors of this special issue, Martin Hilpert and Hubert Cuyckens, and my colleague Belén Méndez-Naya for helpful comments on an earlier version of the article.

Funding: I also gratefully acknowledge the financial support of the European Regional Development Fund and the following institutions: the Spanish Ministry of Economy and Competitiveness (grant FFI2014-52188-P) and the Autonomous Government of Galicia (grant GPC 2014/004).

References

ARCHER 3.1 = A Representative Corpus of Historical English Registers version 3.1. 1990–1993/2002/2007/2010/2013. Compiled under the supervision of Douglas Biber & Edward Finegan at Northern Arizona University, University of Southern California, University of Freiburg, University of Heidelberg, University of Helsinki, Uppsala University, University of Michigan, University of Manchester, Lancaster University, University of Bamberg, University of Zurich, University of Trier, University of Salford, and University of Santiago de Compostela.


HC = *The Helsinki Corpus of English Texts*. 1991. Department of Modern Languages, University of Helsinki. Compiled by Matti Rissanen (Project leader), Merja Kytö (Project secretary); Leena Kahlas-Tarkka, Matti Kilpiö (Old English); Saara Nevanlinna, Irma Taavitsainen (Middle English); Terttu Nevalainen, Helena Raumolin-Brunberg (Early Modern English).


Schneider, Stefan. 2007. *Reduced parenthetical clauses as mitigators*. Amsterdam: John Benjamins.

