Another Heavy Road of Decompositionality: Notes from a Dying Adverb

Remus Gergel
Andreas Blümel
Martin Kopf
Another Heavy Road of Decompositionality: Notes from a Dying Adverb

Abstract
This contribution reports on a pilot case study conducted on historical corpus data from Old and Middle English (primarily Taylor et al. 2003, Kroch and Taylor 2000) and concerned with the ramifications of an ambiguous adverb (eft, ‘again’) at the syntax-semantics interface. The disappearance of the adverb is linked to the development of again’s partially similar functions during the Middle English period. Akin to studies on the adverb again (e.g. Fabricius-Hansen 2001, Gergel and Beck 2015, Beck and Gergel 2015), at least some instances of eft are claimed to require a particular type of analysis of so-called decompositional adverbs, namely one based on lexical rather than only on structural factors. Furthermore, the development shows characteristics of a cyclical development (cf. Jespersen 1917, van Gelderen 2011).
Another Heavy Road of Decompositionality: Notes from a Dying Adverb

Remus Gergel, Andreas Blümel, Martin Kopf

1 Introduction

In this paper we report on a pilot case study conducted on historical corpus data from Old and Middle English (primarily Taylor et al. 2003, Kroch and Taylor 2000) and concerned with the ramifications of an ambiguous adverb (eft, ‘again’) at the syntax-semantics interface. The chronological outcome, which is situated towards the end of the Middle English period in the corpus data, is the disappearance of the adverb and appears to be related to the development of again’s partially similar functions during the Middle English period. We make two suggestions: (i) that similarly to again, there is reason to believe that at least some instances of the adverb eft require a particular type of analysis of so-called decompositional adverbs, namely one based on lexical rather than purely structural factors; (ii) that the development bears the signature of a cyclical development (Jespersen 1917, van Gelderen 2008, 2011, Wallage 2013) in the area of iterative adverbs. The structure of this contribution runs as follows. Drawing on much previous literature (see e.g. von Stechow 1996, Fabricius-Hansen 2001, Beck 2005, Gergel and Beck 2015, Beck and Gergel 2015), we lay out a minimal background on decompositionality for readers unfamiliar with it in Section 2.1 and illustrate the change undergone by again (under inclusion of the issues it poses for one version of the theory) in 2.2 and 2.3. Section 3 lays out the methodological steps undertaken for extracting the meanings of the extinct adverb eft. Section 4 discusses our findings and some of their consequences.

2 Background Decompositionality: The View from Again

2.1 Decompositionality

Decompositional items (sometimes understood in different ways) have played a key role in addressing the syntax-semantics interface due to ambiguities often tied to structure (e.g. von Stechow 1996, Beck 2005). For instance, the adverb again or its German counterpart wieder show an ambiguity between a restitutive and a repetitive reading. Note that adverbs of this type do not necessarily need to decompose themselves at the level of interpretation (as e.g. negative quantifiers or other items may do). What decompositionality in this case refers to is that they can modify meanings of different subparts of the structures they attach to. In the case of again, presuppositions play a key role. They seem to co-vary with the portion of structure that is modified (result state vs. ‘full’ event). The classical ambiguity is exemplified by (1):

1. Leo opened the door again.

As many linguists have pointed out, (1) is two-way ambiguous (Dowty 1979, von Stechow 1995, Beck 2005, Pedersen 2010, among others) between a repetitive reading and a restitutive reading. According to the former, Leo opened the door and that happened before. According to the latter, Leo opened the door and the door had been open before, i.e. Leo’s opening the door does not repeat itself, but rather the door’s state of being open recurs.

Within compositional semantics, one can distinguish two families of analysis to account for this behavior. According to the structural view, ambiguities arise as the result of different syntactic positions of again: one and the same entry of the adverb is used in both contexts. This derives the two distinct readings via some version of a distinction in attachment site. According to this analysis, it is sufficient to say that the single denotation given in (2) below resides either high or low in the VP-structure, delivering a repetitive or restitutive reading respectively. (We refer rather simplistically to ‘VP-structure’, but clearly more options in the extended spine of the VP are compatible with the view; cf. von Stechow’s own original proposal, or e.g. Johnson 2004 for vP.)

2. \[[again_{rep}]] = \lambda P. \lambda e : \exists e’[e’<e \& P(e’)].P(e)
‘This has happened before.’

A lexical ambiguity analysis emerged as an alternative or supplement to this view. Accordingly, there is reason to believe that two distinct entries for again are needed to account for counterdirectional readings, which (2) does not accommodate (e.g. Fabricius-Hansen 2001). Consider (3) and imagine a context where Leo sits down and then acts as described in (3). Next to the semantic aspect that Leo’s state of standing upright is restored, the sentence expresses that Leo’s jumping is the reversal of a previous event, in some cases even a reversal of its direction.

(3) Leo jumped up again.

Again so understood thus contains the presupposition that a counterdirectional event happened before the event modified by the adverb. Thus in addition to the repetitive entry, there is a counterdirectional entry that guarantees this:

(4) \([\text{again}_{\text{ctrdir}}] = \lambda P. \lambda e : \exists e' [e' < e \& P(e')]P(e)\)

‘The reverse has happened before.’

2.2 Recent Change in Again and the Role of Structure

More recently, the behavior of decompositional adverbs like again has been studied diachronically (Beck et al. 2009 for 19th c. again, Gergel and Beck 2015 for Early Modern English). Specifically, Beck et al. (2009) discovered that there are differences between the 19th century use of again and current usage on the basis of corpora consisting of correspondence and similar data (cf. their paper for the composition of the data base). Their main findings are summarized under inclusion of the frequencies obtained in (5) below.

(5) 19th c. English: 21.1% restitutive/counterdirectional again

‘return again’ = come back,
‘connect again’ = put back together

Present-day English: 12.6% restitutive/counterdirectional again

%# return again = come back
%# connect again = put back together

The predicates given in (5) for illustration offer a flavor of the fact that while less transparent predicates such as return or connect were commonly found on restitutive readings in the 19th c. data, they are not accepted by all speakers on such readings in Present-day English. This observation led Beck et al. (2009) to accommodate the change observed between 19th century correspondence data and today’s usage via the following parameters (based on Beck 2005; cf. also Rapp and von Stechow 1999 for a parametric approach to decompositional adverbs):

(6) An adverb can modify
i. only independent syntactic phrases
ii. any phrase with a phonetically overt head
iii. any phrase.

The default setting of the parameters above is assumed to be (i). The idea for the alteration just described is that again is changing (between 19th c. and PDE) from setting (iii) to setting (ii) based on a parametric choice related to structure.

We have used both the term “restitutive” and “counterdirectional” above. This rather clumsy way of referring is motivated by findings in more recent studies of again, concerned with older stages of English, to which we will turn next.

2.3 Diachrony One Step Further and the Lexical Analysis

Early Modern English (ca. 16th and 17th centuries) shows an even higher incidence of coun-
terdirectional/restitutive readings of *again*. Gergel and Beck (2015) analyzed Early Modern English correspondence data (Taylor et al. 2006) and obtained the overall frequency of counterdirectional readings given in (7).

(7) Early Modern English: 41.5% restitutive/counterdirectional

The change between Early vs. Late Modern English as a whole is statistically significant (cf. Gergel and Beck 2015 for details). In addition to the numerical difference, however, an important qualitative distinction in the potentially restitutive readings emerges. Consider the following Early Modern examples of *again*:

(8) and doe looke every oure to **hear from him again**.
(9) …**wryte again** to hym
(10) Tis like people that talk in theire sleep, nothing interupts them but **talking to them again**

Such readings do not have a plausible result state the restitution of which would be referred to in the examples given and hence favor a counterdirectional analysis. By and large, readings of this type cease to be available by the Late Modern English period for *again*.

So far, we have compared Late Modern and Early Modern English on the basis of correspondence data. The development of *again* can, furthermore, be confirmed on a broader scale. Gergel (2012) notes that the ratio of repetitive and restitutive readings of *again* is not stable, or random, when considered over larger periods and in larger historical corpora of mixed genre such as the Penn-Helsinki corpora of historical English (PPCME2, YCOE, PPCEME, PPCMBE). Rather, the proportion of repetitive readings of *again* increases significantly at the expense of restitutive readings (Figure 1 below summarizes this earlier finding).

![Figure 1: Restitutive/counterdirectional and other readings of *again* for ME and ModE.](image)

Both analyses described above are needed to capture the dynamics of change. The structural analysis can handle the change between 19th c. English and PDE as a parameter at the syntax-semantics interface (cf. Rapp and von Stechow 1999, Beck 2005, Beck and Snyder 2001). However, significant data with entirely distinct interpretations from the Early Modern period and earli-
er stages have no plausible result state that could be modified by a unique repetitive/restitutive entry, as examples like (8–10) suggest. Thus the counterdirectional entry (and with it: the lexical analysis of *again*) is required by early English.

Given such quantitative and qualitative results, the question arises: What is a plausible way of modeling the shift between a counterdirectional entry and a structural analysis which can modify result states and genuinely decompose? Beck and Gergel (2015) and Gergel and Beck (2015) suggest adapting a concept originally developed in a synchronic context by Beck (2012) to the issues of language change:

(11) Constant entailments

Variability in the meaning of an expression $\alpha$ between interpretations $\alpha'$ and $\alpha''$ is promoted by the existence of contexts $\phi$ in which an occurrence of $\alpha$ under both interpretations $\alpha'$ and $\alpha''$ leads to the same proposition $\phi'$.

Constant entailments offer a new theoretically-based restriction on the evolution of semantic change within compositional semantics (see, e.g., von Fintel 1995, Eckardt 2006, Deo 2006, Con- doravdi and Deo 2010, Gergel 2009, 2014 for discussions of others). However, unlike some of the received wisdom filtered from the more widely documented cases of traditional grammaticalization, the concept by and of itself does not require a particular type of a priori trajectory or directionality. Hence, interesting empirical questions arise. First, how do other potential decompositionists evolve? For example, some of the developmental steps reported may well be conditioned by *again*’s original morphosyntax and semantics, while others may not be tied to it but rather more general in nature. (Compare German *wieder*, ‘again’, Fabricius-Hansen 2001, for a potentially similar trajectory to the one undergone by *again*, though more corpus studies may offer further insight in this area as well.) Second, what are the effects when two such markers are available in one and the same language and interact diachronically? Given that repetitivity, counterdirectionality, and similar concepts can be expressed in a multitude of ways, this desideratum looms large too, clearly so for the markers that seem to be grammaticalized solidly with such functions. Our aim is not to set up definite general answers to these questions here, but our study has been aimed at doing some of the necessary beginning work towards opening up the possibility of such comparisons on the basis of corpus work. Before delving into comparisons and possibilities for routes of development, we will offer an overview of the methodological issues encountered in the analysis of the investigated adverb *eft* next.

3 Methods for the Present Study

3.1 Data Selection

The present claims are based on the detailed evaluation of ca. 850 tokens of *eft* together with their contexts. At the center of the current work are the PPCME2 and YCOE corpora of Middle and Old English, respectively, Kroch and Taylor (2000), Taylor et al. (2003). The YCOE contains by far more instances of *eft* than the PPCME2, both in absolute terms and in terms of frequencies per tokens. In the chronological period subsequent to Middle English, i.e. Early Modern English, on the other hand, *eft* was not productively available any longer. The large PPCEME corpus (Kroch et al. 2006) only has two tokens and it was discarded for our study. We searched the following basic adverbial forms of *eft* in the corpora (in the standard notation of the parsed corpora):

(12) ef, Ef, eft, Eft, efte, +afte, +Afte, heft, Heft, eft+d, Eft+d, eaft, Eaft, efft, Efft

Given the overall corpus situation described, we have exhaustively extracted the occurrences of *eft* and its variant simplex forms (i.e. no morphological composite forms) from the Middle English subperiods of the PPCME2 and matched them against a sample consisting of a somewhat larger number of occurrences of the same set of forms in Old English texts from the YCOE. The restriction on the Old English data was made to keep the samples closer in size and based on practical considerations. The selection from the broader OE data populations was conducted by using the randomization function in the software package R. Thus, from each of the abundantly repre-
sented subperiods O2 and O3 of Old English, 200 randomly chosen tokens were selected and annotated. This is a small sample compared to modern studies (but still a comparatively large one for early diachronic semantic studies conducted on the basis of case-by-case contextual annotation of tokens). The randomization procedure ensured that, on simple probabilistic grounds, texts that are widely represented in the corpus population were also similarly widely represented in the samples while sticking to the standard corpus periodization. However, not all historical subperiods posed this kind ‘luxury problem’. The reversed issue of not having a sufficient amount of data for detecting descriptive trends quantitatively also arose. For the marginal periods within the span in which *eft* was investigated, i.e. the initial subperiod O1 of Old English and the final subperiod M4 of Middle English, this did not matter immediately for discerning the inner dynamics of the developments. Nonetheless it has to be kept in mind that we are looking—in terms of more reliable quantitative developments—at a truncated time axis. While some form of truncation (whether we are aware of it or not) is almost always given in historical linguistics, due to the familiar actuation problem of change, it has to be hoped that in further work more extended samples can be taken into consideration. Given the provisos, the six contiguous corpus subperiods we considered are as follows:


Within the non-marginal subperiods of the parsed corpora, O4 and M2 posed a quantitative issue due to their low incidence of *eft*. We have therefore annotated data from additional sources that fall strictly within the two periods, respectively.

Finally, the parsed corpora also contain texts with ‘mixed’ labels, in which composition and manuscript date are assumed to diverge. For instance, M23 indicates that a text was composed in the second Middle English period, but the manuscript dates from the third one. The mixed periods contained very few hits in the case of *eft* and we will not be concerned with them here for the most part. The only notable exception was MX1 (where X stands for undetermined composition date), which has therefore been considered and added to the picture. It can be a window into the connection between Old and Middle English, but given the unknown composition date of the texts its evaluation must also be considered with caution.

### 3.2 Semantic Annotation

The categorization of the tokens fell into three major classes. As with *again*, there were repetitive and restitutive readings of *eft* (the latter are also known as ‘counterdirectional’—cf. Fabricius-Hansen 2001 and Sections 2 and 4—but for first descriptive purposes, this did not matter). A sufficient condition for a restitutive or counterdirectional reading, the salient presence in the preceding context of a state that is restituted or, more frequently in practice: of a counterdirectional event. Repetitive readings required a fulfilled presupposition of the entire event. Furthermore, we have taken into account temporal readings of *eft*, which do not normally appear with *again*. They can be translated as ‘then’ or ‘afterwards’ and typically appear in narrative sequences; crucially, they do not involve presuppositions of repeated or reversed events.

In a few cases, general knowledge (e.g. about religious or historical narratives) was considered, but in the overwhelming majority of cases, then, the label with which an occurrence was categorized with respect to the three major readings was decided primarily on the basis of the local context. To summarize, we considered the following: (i) did the event hold before; (ii) was only the result state positively given in preceding context and is (usually after a counterdirectional event) the state restored (cf. pairs in ModE *turn off*/*turn (back) on again*); (iii) is there a narrative sequence which makes *eft* (often preposed) mostly compatible with a temporal reading, ‘afterwards’ or ‘then’? The following examples illustrate the three readings:

(14) *Eafterward me ssel þerne mete eft c hyewe /a se þe oxe þet...*  
    *afterward one shall this food again chew/as the ox that...*
    *(CMA YENBI111.2146) [repetitive reading]*

(15) *ðe feordæ t ime w es ðoa ha misde hire sune & eft him ifunde.*
the fourth time was at-that-when she missed her son and again him found.
(CMANCRW1,II,62.651) [restitutive /counterdirectional]

(16) *Eft* ða þa Iulianus ... ðearð ðas to casere gecoren,...
afterward when Julianus was to emperor chosen
(coaelive,+ALS[Agnes]:394.1990) [temporal reading]

In cases in which only one reading could be ruled out, we noted any of the three possible pairwise ambiguities. We aimed to establish a clear choice between counterdirectional and repetitive readings whenever the context allowed for it. In the vast majority of cases this turned out to be the case. As in other studies on decompositionalty, some examples remained ambiguous with regard to the repetitive vs. counterdirectional distinction after cross-checking; their labeling then was marked as such. The temporal readings could frequently not be ruled out. We hence chose to keep track of them as a latent meaning (when it was available) even if the most salient reading was clearly repetitive or counterdirectional (for reasons of space we will not discuss these numbers here). A disjunctive forced choice of temporal readings as delimiting them from the other crucial readings was not useful in the majority of cases of this study. We have made a distinction, however, between instances of *eft* that can only be understood as ordering temporal sequences (these are reflected in the figures below, cf. Figures 2 and 3) and items that have temporality as meaning that is possible in addition (to repetitive or counterdirectional/restitutive readings). A few examples had to remain entirely unclear. Philological translations were consulted in the extant cases, but we documented our decisions on the basis of the primacy of the contexts studied in each individual case; translations could not always be ‘followed’ (e.g. they crucially do not disambiguate simply by translating *eft* via ‘again’). Finally, notice that *eft* can have further, discursive functions that often join the three main readings described. We leave the possible discourse functions of *eft* aside in this study and hope that they will receive subsequent scrutiny in future work.

4 Findings

4.1 Occurrences of *eft* and the Distribution of Counterdirectional/Restitutive Readings

As noted in Section 3, the adverb *eft* was available over two major stages of the language, Old and Middle English. The incidence of restitutive/counterdirectional readings in the data is as in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>tokens</th>
<th>restitutive/cd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old English</td>
<td>488</td>
<td>115</td>
</tr>
<tr>
<td>Middle English</td>
<td>362</td>
<td>123</td>
</tr>
</tbody>
</table>

Table 1: Incidence of restitutive/counterdirectional *eft*.

While there is a clear descriptive rise in the incidence of restitutive readings in the Middle English period as a whole, this tendency is less transparent to analytical interpretation than it may seem. Recall that our counting is based on data analyzed from the subperiods O2, O3, O4, MX1, M1, M2, and M3. Consider Tables 2 and 3, as well as Figure 2.
Two observations emerge from the development over the subperiods. First, while Middle English ranks higher in terms of averaged restitutive/counterdirectional readings than Old English, as we have seen, there is no continuous rising tendency to be discerned over its major subperiods. And second, while there is a clear peak of restitutive readings reached during the subperiods MX1/M1, a falling tendency can be observed afterwards. We will return to these points in the discussion of Section 4.2.

A more general observation in terms of the quality of the data containing the adverb *eft* is that in both Old and Middle English, genuine counterdirectional readings—often accompanied by actually given counterdirectional events in the context—are available. We give an exemplary selection of the predicates obtained with counterdirectional occurrences of *eft* for the subperiods O2 and MX1 in (17) and (18) below. Motion predicates such as ‘come’ or ‘go’ recur several times in the data, but a broad range of predicates can generally be observed.
(17) Predicates with counterdirectional readings of _eft_—period O2


(18) Predicates with counterdirectional readings of _eft_—period MX1


4.2 Discussion

_Eft_ shows certain parallels with _again_ as far as the basic ambiguity we started out with and the availability of counterdirectional meanings at early stages are concerned, but a series of differences become apparent. That is the case even beyond the availability of the third temporal reading. Given the oscillations in the data, we suggest that thinking of possible developments in terms of three time spans may be useful. One is the Old English period (we do not go into any major subdivisions of it in the discussion). A second one is the early Middle English period consisting of MX1 and M1 and finally, the third one is the final stretch in the visible life of _eft_ within our data, i.e. M2 and M3.

Crucially, we could observe no clear numerical development, say with a steady increase or decrease of restitutive/counterdirectional readings throughout the subperiods studied, even when the data are clustered as just described. On a null hypothesis, this may seem unspectacular. Perhaps nothing of crucial importance (as far as the grammatical factors were concerned) happened to _eft_ at all and all we see in the course of its lifetime is sheer noise. But there are two major factors that lead us to believe otherwise. One is given by the interaction of _eft_ with _again_, the other by a prediction of the notion of constant entailments. We discuss issues relating to the interaction of the two adverbs below.

_Again_ was originally a preposition. It begins to be visible as an adverb in the data of the parsed corpora during the Middle English period. However, it does so rather sporadically at the beginning of the period (cf. Gergel 2012 and Beck and Gergel 2015 for discussions and updates). That is, whatever was happening to _eft_ at the beginning of Middle English cannot be blamed to a large degree on the appearance of a putative strong competitor on a quantitative basis. By contrast, the development of _again_ in later Middle English may well have influenced the fate of _eft_. There may be possible provisos for the poorly represented M2 period in the parsed corpus data, but the later development of _again_ makes a similar point: restitutive/counterdirectional readings decrease during later Middle English. Another observation is that the incidence of repetitive readings for _eft_ in the cluster M2+M3 as a whole compared to M1 increases. We suggest that at least the decrease in restitutive readings in later Middle English (i.e. the M2+M3 period) is expected due to the interaction with _again_. While the proportion of repetitive readings of _again_ increases somewhat in M3, its overall frequency, i.e. its density per number of tokens increases considerably during the M3 subperiod compared to early Middle English. Hence _again_’s impact as a competitor of _eft_—more specifically, as a competitor with predominantly restitutive/counterdirectional readings—is expected to be reflected more strongly during the later periods of Middle English, when _again_ is several times more frequently available. This interaction may explain at least the fall in restitutive readings of _eft_ in M3.

Finally, we note that the rise in the incidence of restitutive/counterdirectional readings during Middle English as a whole and in particular the periods MX1 and M1 remains a puzzle compared to Old English. We suspect that the role of constant entailments may be used to account for this but we will leave working out the solution to this puzzle for another occasion. For now, suffice it to note that there is no prediction of directionality that emerges from the notion of constant entailments.

A final note is in order regarding the field of linguistic cycles (cf. van Gelderen 2008, 2011, for recent overviews). While _eft_ had originally fulfilled functions taken over by _again_ at later
stages, there was also an extended time at which the two items interacted (similarly e.g. to \textit{ne+not} in English; cf. Fischer et al. 2000 for a succinct description of the complex syntactic facts in the history of English). A preliminary overview of the co-occurrence of the two items is given in the following table:

<table>
<thead>
<tr>
<th>O2</th>
<th>O3</th>
<th>O4</th>
<th>MX1</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4/202=1.9%</td>
<td>7/84=8.3%</td>
<td>4/78=5.1%</td>
<td>4/145=2.7%</td>
<td>4/83=4.8%</td>
<td>3/56=5.7%</td>
</tr>
</tbody>
</table>

Table 4: Co-occurrence \textit{eft+again}.

From the time course of change considered, \textit{again} seems to be a clear contributor to the final demise of \textit{eft}. But two distinctions from the classical case of the negative cycle can be discerned. First, even though interaction existed for a considerable time, co-occurrence was never particularly frequent. The other difference is that unlike in the case of negation (cf. Wallage 2013), it cannot be claimed that the final stages of \textit{eft} should have survived essentially piggy-backing on the more recent item \textit{again} in instances of co-occurrence (while \textit{ne} might have done so with respect to \textit{not}). The co-occurrence does not become more frequent at the final stage M3. More generally, we suspect that there may be two types of developments that appear to be cyclic - the ones in which the two items truly reinforce each other and the ones in which they primarily compete for similar functions (without necessary frequent co-occurrence).

4.3 Summary

The overall historical change undergone by \textit{eft} does not follow an identical numerical or semantic trajectory as \textit{again}. The early adverb supports, however, a counterdirectional analysis of a further potentially decompositional item. In this paper, we have offered basic descriptions and suggested an account of its decrease on restitutive readings towards the end of the Middle English period due to the natural competition with the rising of the adverb \textit{again} with which it produces the appearance of a cyclical pattern from a diachronic vantage point. The increase of restitutive readings in the texts of the early Middle English period may be due to the role of constant entailments, though the specifics of this interaction still require further elaboration.

References


Remus Gergel and Martin Kopf
Department of English / English linguistics, Fachrichtung 4.3.
University of Graz / Saarland University
8010 Graz, AT / 66123 Saarbrücken, DE
remus.gergel@uni-graz.at
martin.kopf@uni-graz.at

Andreas Blümel
Department of German
University of Göttingen
37073 Göttingen, DE
andreas.bluemel@phil.uni-goettingen.de