Explorations in English Historical Syntax

Edited by

Hubert Cuyckens

Hendrik De Smet

Liesbet Heyvaert

Charlotte Maekelberghe

JOHN BENJAMINS PUBLISHING COMPANY

Explorations in English Historical Syntax

Studies in Language Companion Series (SLCS)

This series has been established as a companion series to the periodical *Studies in Language*.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/slcs

Founding Editor

Werner Abraham
University of Vienna / University of Munich

Editors

Werner Abraham
University of Vienna / University of Munich

Elly van Gelderen
Arizona State University

Editorial Board

Bernard Comrie
University of California, Santa Barbara

William Croft
University of New Mexico

Östen Dahl
University of Stockholm

Gerrit J. Dimmendaal
University of Cologne

Ekkehard König
Free University of Berlin

Christian Lehmann
University of Erfurt

Elisabeth Leiss
University of Munich

Marianne Mithun
University of California, Santa Barbara

Heiko Narrog
Tohuku University

Johanna L. Wood
University of Aarhus

Debra Ziegeler
University of Paris III

Volume 198

Explorations in English Historical Syntax
Edited by Hubert Cuyckens, Hendrik De Smet, Liesbet Heyvaert
and Charlotte Maekelberghe

Explorations in English Historical Syntax

Edited by

Hubert Cuyckens
Hendrik De Smet
Liesbet Heyvaert
Charlotte Maekelberghe
KU Leuven (University of Leuven)

Table of contents

Foreword

The chapters in the present volume spring forth from work that was first presented at the 18th *International Conference of English Historical Linguistics* (ICEHL18), held at the University of Leuven, 14–18 July, 2014. ICEHL, which is held in Europe biennially, has alternated since 1992 with the International Conference of Historical Linguistics. Its long tradition goes back to 1979, when it was organized for the first time by the University of Durham. Yet it had never been organized in Belgium before. It is no coincidence that the University of Leuven should be the first Belgian university to have done so. After all, historical linguistics of English has been at the heart of its linguistics department ever since Xavier Dekeyser started a History of English class in 1971.

For turning the conference into a success, we are first of all obliged to all the presenters and conference participants, who created a stimulating atmosphere of intellectual exchange. Special thanks also go to the plenary speakers Charles Boberg, Robert Fulk, Peter Grund, María José López-Couso, and Marit Westergaard, who demonstrated that English historical linguistics is truly a multifaceted research domain. We would furthermore like to thank our co-organizers Peter Petré, Frauke D'hoedt, Lauren Fonteyn, and Nikki van de Pol, and all our student helpers, as well as Lieselotte Brems (Université de Liège). For financial support, we are grateful to the Research Foundation Flanders (FWO).

There is a longstanding tradition of publishing a selection of ICEHL papers with John Benjamins, and we are happy to subscribe to this tradition. To ensure thematic coherence, accepted submissions were divided into two volumes. The present one, which is published in the time-honored *Studies in Language Companion Series*, focuses on syntactic change. A companion volume entitled *Sociocultural dimensions of lexis and text in the history of English* is to appear in another venerable series by Benjamins, *Current Issues in Linguistic Theory*. For the present volume we have made a careful selection from a generous number of submissions. This would not have been possible without all of the detailed review reports, for which we would like to thank all the reviewers: Karin Aijmer, Kristin Bech, Laurel Brinton, Timothy Colleman, Bridget Drinka, Matthias Eitelmann, Christine Elsweiler, Þórhallur Eyþórsson, Olga Fischer, Victorina Gonzalez-Diaz, Yoko Iyeiri, Kristin Killie, Merja Kytö, Bettelou Los, Brian Lowrey, Christian Mair, Belén Méndez-Naya, Ruth Mohlig-Falke, Arja Nurmi, Günter Rohdenburg, Juhani Rudanko, Anni Sairio, Hans-Jörg Schmid,

Exploring English historical syntax

Hubert Cuyckens
KU Leuven

1.   A renewed appreciation for English historical syntax

In the past thirty years or so, the study of English historical syntax has met with increasing interest (witness two monographs that were published just recently: *A Historical Syntax of English* (2015) by Bettelou Los and *A Brief History of English Syntax* (2017) by Olga Fischer, Hendrik De Smet & Wim van der Wurff). In these years, we have not only seen a wide variety of topics being discussed (see, e.g., Fischer et al. 2017: 4–6 for an overview of syntactic patterns and their changes), but these topics have been addressed from several angles and theoretical frameworks (the philological tradition, the generative framework, cognitive-functional and usage-based approaches, sociolinguistic approaches). Accordingly, more information, data, and insights on the syntactic history of English have become available, which, in turn, have served as a basis for linguists to formulate hypotheses about, and explore theoretical issues in, syntactic change and language change in general.[1]

Linguists' current interest in English historical syntax is not an isolated phenomenon. Aside from the fact that work on the history of English syntax is likely to generally benefit from English being the most widely studied language (see Brinton & Bergs 2017: 1),[2] it currently also benefits from the fact that the field of (English) historical linguistics has seen something of a renaissance. While linguistics in the first half of the twentieth century (structuralism, generative linguistics) was not

---

1.   Note that Trousdale (2017) reserves the term "syntactic history of English" for "particular development[s] in a particular set of varieties, based on a particular set of evidence" (2017: 103), and the term "English historical syntax" for general issues in syntactic change, as they are informed by the data from the history of English.

2.   In this respect, they state that "… it seems safe to say that the majority of studies in contemporary linguistics [both synchronic and diachronic] deal at least in part with English" (Brinton & Bergs 2017: 1).

exactly accommodating of the study of (English) language change, a renewed appreciation can be observed in recent years (see also Bowern & Evans 2015: 1). Testimony to that are the six-volume *Cambridge History of the English Language* (1992–2001), edited by Richard Hogg, the *Oxford History of English*, edited by Linda Mugglestone (2006), several handbooks of English historical linguistics published by major publishers (van Kemenade & Los 2006; Bergs & Brinton 2012; Nevailainen & Traugott 2012; Kytö & Pahta 2016), and two recent textbooks (Brinton 2017; Brinton & Bergs 2017). Historical linguistics in general has also seen the publication of several handbooks (Joseph & Janda 2003; Bowern & Evans 2015) as well as the establishment of new journals (*Journal of Language Dynamics and Change* (Brill); *Journal of Historical Linguistics* (John Benjamins); *Journal of Historical Syntax* (online section of *Language*)).

In accounting for this strengthened interest in English historical linguistics, and English historical syntax in particular, we would like to single out two factors. First, the creation of digital text corpora for the study of older language stages and language change has given researchers access to large quantities of data. This has not only greatly facilitated historical work, but has also opened up new research topics and has led to new methodologies. There can be little doubt that English has been one of the languages whose study benefited most from this development. Second, since the 1980s, the frequency and breadth of studies in (English) historical linguistics have greatly expanded as appreciation grew for variation and change as an integral part of language. The linguistic approaches/frameworks that can be said to foster (English) historical linguistic work most easily are the cognitive-functional, usage-based, and sociolinguistic approaches to languages: they base their analyses on performance-based data, viewing "performance variants" in particular as "central to language change (Fischer et al. 2017: 30). Then again, generative grammar as well is increasingly integrating diachronic work (on English), working on the assumption that "language change can provide insights into the (generative) language faculty" (van Gelderen 2016: 81). In the following paragraphs, these two factors will be discussed in more detail. Note that while undoubtedly each factor will have contributed individually to the increased appeal of English historical linguistics/syntax, they will also have reinforced each other. (Digital) text data potentially holding evidence of variation will have likely benefited performance-oriented diachronic analyses that view variation and change as an integral part of language, as well as competence-oriented models interested in our internal grammars. Conversely, linguists' orientation towards historical (syntactic) data will have stimulated the compilation of new and richer corpora.

As Kytö and Pahta (2016: 1) point out, "Research in English historical linguistics is firmly anchored in evidence drawn from texts." Until the 1980s, these were almost exclusively "texts on paper" (Fischer et al. 2017: 8). And while the use of paper texts

has a long tradition and work emanating from this textual material has been inspiring, "there are obvious problems that it faces due to the nature of the data source" (Fischer et al. 2017: 11). An important problem is that reading texts to find examples of particular syntactic phenomena can be very time-consuming and that it is nearly impossible to read all texts of a particular time period. Accordingly, data collection is likely to be incomplete and analyses based on these data may be impressionistic or may not be replicable.[3] At the same time, the careful reading of texts "ensures that the researcher gets a sense of the complete syntactic system of the language of the text" or that "expectations based on present-day English will … colour the perception of earlier data" (Fischer et al. 2017: 11–12). Since the 1980s, paper texts as a data source for work in (English) historical linguistics/syntax have gradually given way to corpora, structured collections of texts typically available in digital format. After the publication in 1991 of the *Helsinki Corpus*, "the mother of all diachronic English corpora" (Fischer et al. 2017: 12), we have seen a robust growth of corpora for the diachronic study of English linguistic material, such that "the amount and variety of historical corpora available for English today is much larger than for any other language" (López-Couso 2016: 129). A listing of the most relevant corpora can be found in Brinton & Bergs (2017: 2) and in Fischer et al. (2017: 12–14); a detailed discussion of these corpora is presented in López-Couso (2016).

Obviously, data collection and analysis making use of electronic (diachronic) corpora is far less time consuming than by reading through texts on paper; it is also a better guarantee of full accountability to, and replicability of, the data (see Fischer et al. 2017: 14). In addition, diachronic electronic corpora have opened up new possibilities for the scholar of English historical syntax. As electronic corpora have come to include sociohistorical markup and specifications for genre, register, and variety, the study of external factors in linguistic/syntactic change has become an increasingly popular research topic (next to the study of formal and functional changes). Methodological advances such as the development of search software have greatly facilitated the study of less frequent and/or more complex syntactic structures, and statistical tools have been introduced to gauge the significance of changes in (syntactic) phenomena. Finally, the greater diachronic data sets generated by large corpora have given rise to research questions that go beyond the mere description of change, but that probe into the internal dynamics of (syntactic) changes; as Hilpert and Gries (2016: 38) indicate, this means that statistically informed approaches are required that are multifactorial, that involve interactions between formal, functional, and social predictors, and that involve interactions with time.

---

3. "Up until quite recently, listing and classifying data drawn from often haphazardly collected datasets was considered sufficient, and conclusions were drawn on impressionistic generalizations based on the data" (Kytö & Pahta 2016: 1).

It should be kept in mind that using corpora as data sources may also have disadvantages. Fischer et al. (2017: 14) identify the following (based on work by Viana et al. 2011): (i) what is easily retrieved through corpus searches is not necessarily what is linguistically most relevant; (ii) an exclusively quantitative approach may be too reductive; (iii) substantial amounts of data may lead to unwarranted generalizations if it is insufficiently recognized by researchers that the data only offer a partial picture of historical reality; (iv) caution should be exercised when collecting and comparing data from various corpora to achieve greater representativeness.

A second factor that may account for the increasing interest in English historical syntax is the rise of linguistic frameworks – cognitive-functional linguistics, usage-based linguistics, and sociolinguistics – that give sufficient attention to change. Their approach to change is insightfully described in Fischer et al. (2017: Chapter 3), on which the brief account that follows is based. Generally, what these frameworks share is their interest in language variation and the view that "variation is the seed of change" (Fischer et al. 2017: 29). Central to language change are, in particular, the "performance variants, to be found in the output of adults" (30). This variation at performance level comprises variants of linguistic structures that may result from analogy with other structures, from pragmatic inferencing, or from contact with speakers of other languages. In addition, variation in linguistic structure may correlate with such performance-related phenomena as the frequency of the structure or the register/genre in which the structure occurs. In all, these performance variants can be said to result from language-internal as well as language-external factors. Importantly, in these performance-based approaches, the makeup of our mental grammars, and our knowledge of changing grammars, can only be obtained indirectly, by studying language output. In historical linguistics, this output obviously consists largely of written text (nowadays mostly in digital format).[4]

Linguists' increased interest in language change since the 1980s (in approaches that view variation and change as an integral part of language) is not a new phenomenon, as interest in the history of a language had also characterized the nineteenth-century philological approach to language. With these nineteenth-century philologists, current historical linguists also share a reliance on written texts, a focus on language output and language use, and an attentiveness to external circumstances (the historical, sociocultural background).[5]

---

4.   See Mair (2016), however, for a survey of oral historical data.

5.   As Fischer (2007: 59) states, "[philologists'] interest was intimately connected with literature, history, and culture of national communities…: language forms were described in close connection with language use; form was bound up with function/meaning."

Linguists' renewed appreciation of language change since the 1980s is, however, less easily reconciled with the structuralist movement that started at the beginning of the twentieth century and with the views of the early generative linguists in the 1960s and 1970s. Structuralists' and generative grammarians' focus was on laying bare the system behind/underlying language (*langue*, language competence, I-language). Accordingly, structuralists paid more attention to *langue* than to *parole* 'performance', "the study of language utterances as they are processed by speakers and hearers against a historical, sociocultural background" (Fischer et al. 2017: 29). This shift of interest led to a diminished study of the historical development of languages, which relies precisely on performance-related output (in the form of texts) produced in a particular sociohistorical setting. On the structuralist approach, then, "change could only be detected *indirectly* via a comparison of *synchronic* linguistic states or systems" (Fischer et al. 2017: 29). For early generative grammarians, language change was not a research priority either. In their view, language competence or internalized grammar could only be detected by introspection, that is, by probing one's own native speaker grammatical knowledge or by asking other native speakers for grammaticality judgments. Because native speakers of the historical stages of a language (e.g. Old, Middle, Modern English) are not available, interest went down here as well.

While van Gelderen (2016: 83) still calls the relationship between historical and generative linguists "uneasy", she points out that there have always been generative linguists interested in language change, as change – more so than language comparison – gives insight into the smallest variations Universal Grammar will tolerate. For instance, a frequently attested change such as that from preposition to complementizer, and not the other way around, is "a unidirectional change that gives us insight into the language faculty" (van Gelderen 2016: 81). Testifying to generative linguists' interest in language change are, for instance, the DiGS (Diachronic Generative Syntax) conferences that started in 1990. Many scholars of generative syntax have also embraced work with the *Penn Corpora of Historical English* (see van Gelderen 2016: 83). Well-researched topics in the generative historical syntax of English are, for instance, the change from Object – Verb to Verb – Object word order (Canale 1978; van Kemenade 1987), the development of the modals (Lightfoot 1979, 1991, 1999), and more recently the relation between word order and information structure (van Kemenade 2009; Pintzuk & Taylor 2011; Los 2012). Importantly, for generative historical linguists, the proper object of investigation is not so much "language change" but "grammar change", with grammar change occurring as a result of reanalysis during acquisition (see van Gelderen 2016: 328). As van Gelderen (2016: 81) puts it:

> Generative grammar is interested in how a child acquires a grammar on the basis of the language the child is exposed to. If the language the child hears has changed or is changing from that which the parents/caregivers grew up with, the child will have a different input and may come up with a grammar (I-language) different from that of the preceding generation. … External change modifies the linguistic input, the E-language available to the child, and the real interest is how the child deals with this in terms of parameter resetting. If children hear more Verb – Object (VO) sentences than Object – Verb (OV) sentences, … they will assume that the word order is VO and set their parameter for Verb-headedness as head-initial rather than as head-final.

The performance-oriented, broadly functional and the competence-oriented, generative approaches to language change (and historical syntax in particular) are described in great detail in Fischer et al. (2017: 31–41). Suffice it here to state that the broadly functional approach focuses on language output, considers language use as the locus of change, and views change as gradual. In turn, the generative approach focuses on the language system, considers language acquisition as the locus of change, and views change as radical.

2. Factors of change

With regard to the factors conditioning syntactic change, the literature usually agrees that the language-internal factors *analogy* and *reanalysis*, as well as the language-external factor *contact* are at play (see, e.g., Harris & Campbell 1995). In addition, *frequency* has been seen to play a role (Bybee 2007; Hilpert 2013). Additional external factors that have been proposed are of a *social* nature or should be situated in the realm of *text-type* and *register*. It should be noted that in the generative approach to syntactic change, reanalysis during acquisition is the primary factor of change (involving parameter resetting); the broadly functional approaches, however, envisage a mix of internal and external factors in accounting for syntactic change. The following paragraphs can only present a brief discussion of factors of change; a general survey of factors of change can be found in Fischer (2007: 126).

A core factor that is discussed in the literature on syntactic change is reanalysis (or neoanalyis; see Andersen 2001; Traugott & Trousdale 2013). The foundational definition is given by Langacker (1977: 58): "change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation". An often quoted example is the *for…to*-complement construction, which arguably has been reanalyzed from a construction in which *for* + NP belonged to the main clause. An example provided by Harris & Campbell (1995: 62) is the reanalysis from (1) to (2), as can be seen from (3), where the *for… to* constituent has been pre-posed:

(1)   [*it is bet for me*] [*to sleen my self than ben defouled thus*]
      'It is better for me to slay myself than to be violated thus.'

(2)   [It is better] [for me to slay myself than to be violated thus.]

(3)   [For me to slay myself] [would be better than to be violated thus.]

Another well-known case is the shift from [*a lot*]$_{NP}$ + partitive genitive, as in [*a lot*][*of land*] to [*a lot of*]$_{Quantifier}$ + N, as in [*a lot of*][*land*]. The new surface pattern becomes visible as soon as the quantifier syntactically behaves like the quantifiers *much/many*, whereby verb agreement occurs with N (which could be singular or plural) rather than with singular *lot* (see Brems 2010). In contrast to reanalysis, analogy does not involve modification of a surface manifestation. Rather, analogy refers to "a condition where a … similarity exists between two (or more) items, or classes, or constructions, etc." (Harris & Campbell 1995: 51). The suggestion is often made that reanalysis is primary and analogy secondary (see, e.g., Hopper & Traugott 2003); however, Fischer et al. (2017: 47) argue that "the *perception of similarity* must be logically primary to the reanalysis" (see also Fischer 2007; De Smet 2009, 2013). A prime example of analogy-based change in the history of English syntax is the detailed discussion of the diffusion of the gerund in De Smet (2013). Another well-known example is the analogy-based extension of the *be going to*-future (see, e.g., Hopper & Traugott 2003: 2–3; Fischer et al. 2017: 46–47).

  With regard to contact and borrowing as external factors in syntactic change, we would like to refer to the detailed discussions in Fischer (2013) and Fischer et al. (2017: Chapter 4). Social factors have also been argued to foster syntactic change, whereby social norms or practices increase the discourse frequency of particular structures (Simpson 2002). Several instances can be found in Nevalainen & Raumolin-Brunberg's (2003) monograph on *Historical Sociolinguistics*. A fairly recent example is the increasing use of the gender-neutral use of anaphoric *they/them/their* instead of *he/him/his* and *she/her* as a socially determined avoidance strategy (Stein 1990: 329). Providing additional insight into syntactic change is the variation in text-type or register. In this respect, Foster and van der Wurff (1995) argue that prose is the locus of change of the shift from pre-verbal to post-verbal objects: in the 14th and 15th centuries, "poetry and prose increasingly diverge in their use of pre-verbal objects, with prose showing a clear decline but poetry maintaining stable levels" (Fischer et al. 2017: 17). Another example is the increasing use of *with*-augmented absolute constructions, as in (4), from Early Modern English to Present-day English; here it is spoken language that promotes the newer, augmented variant (see van de Pol & Cuyckens 2014; van de Pol 2016).

(4)   **With** Stannis Baratheon and Tyrion Lannister contending for the iron Throne, we have a rare chance to improve our lot.
                              (George R.R. Martin. 2005. *A Feast for Crows*, p. 185)

A final factor often mentioned in the literature is frequency. In addition to frequency correlating with erosion (especially in grammaticalizing items) (see, e.g., Bybee 2003, 2007), frequent patterns may also function as analogical attractors (i.e. as patterns for other forms). Fischer et al. (2017: 48) provide the example of word-order patterning: "Thus in English, the SVO word order, which already in OE was more frequent than other possible orders like SOV and VSO, became more or less the only available word order in Modern English". In D'hoedt & Cuyckens (2017), it is shown that in the development of the Secondary Predicate Construction introduced by *as* (i.e. the *as*-SPC; see (5)), the zero-SPC (6) functioned as the more frequent model construction, whose distribution was to a large extent copied by the *as*-SPC.

(5)   Under the notion of Protestants, we should consider ourselves **as** christians reformed. (as-SPC)
        (PPCMBE, 1762, J. Burton, *Two Sermons Preached at Saint Mary's*)

(6)   We consider prompt action Ø invariably better than quiet reflection. (zero-SPC)
        (*PPCMBE*, 1908, A.C. Benson, *The Schoolmaster*)

The factors mentioned above cannot all account for the changes in English syntax: a structural change may be affected by other changes or other structural phenomena. As such, word-order change has been seen to be involved in several structural changes such as the rise of the operator *do*, the loss of impersonal verbs, or the increasing loss of the complementizer *that* (see Fischer et al. 2017). In addition, the word-order change from OV to VO has been associated with the rise of the *for… to* infinitival pattern (Fischer 1988; De Smet 2009; Fischer et al. 2017: 204–207), and the increasingly rigid SV order in English may have led to subjects accommodating "whatever information was contextually given" (Fischer et al. 2017: 207), also termed the "permissiveness" of English subjects (see Los & Dreschler 2012).

3.   Contributions to this volume

Against this background, we will now discuss the contributions to the present volume. Obviously, they only offer a selection of the wide range of research topics in English historical syntax.[6] What these contributions in our volume share is that they make use of large sets of text data (obviously, given the historical nature of the data,

---

6.   Syntactic changes in English are numerous and cover a wide range of domains; among them are word-order changes at the level of the NP and the clause, modality, complementation, negation. For extensive surveys of changes in English syntax, see Fischer & Van der Wurff (2006), Los (2015), Fischer et al. (2017).

introspective native speaker intuitions would not be available here). In keeping with the latest developments, these data are largely in digital format (corpora). All papers should be situated within a broadly functional approach, which also pays sufficient attention to the sociohistorical context of the language phenomena. Their object of research is changes in language output rather than changing processes in the underlying grammar system; it is product- rather than process-oriented (see Fischer et al. 2017: 32). Any generalizations made about language users' changing mental/internal grammars need to be seen as interpretations of these data. Interestingly, the fact that none of the contributions strictly fits within the domain of generative grammar supports van Gelderen's (2016: 83) statement that at ICEHL conferences, generative grammar is a minor framework.

The first set of papers in this volume address what is one of the most basic and pervasive tasks of syntax, viz. argument structure. Here, Bettelou Los discusses the rise of "permissive subjects", a trend which, as subjects came to be fixed in initial position, increasingly allowed for contextually given information to be accommodated in the subject slot. From subjects we move on to objects, with Nikolaos Lavidas' study examining cognate noun constructions in Tyndale's new testament. Argument structure naturally relates to clausal complementation. This relation is pursued in Günter Rohdenburg's paper, which compares the evolution of simple complements (prepositional objects and adjuncts) with that of complex verbal complements (*that*-clauses). Also in the domain of complementation, Brian Lowrey examines finite causative complements in Middle English (V+*that* vs. V+NP+*that*), while Yoko Iyeiri considers the variability between non-finite causative *to-* and bare infinitives in Early Modern English, and Mark Davies & Jong-Bok Kim consider non-finite *into*-causatives in American English. Going beyond the core argument slots, a paper that deals with non-finite structures outside the realm of verb complementation is that by Carla Bouzada Jabois, who looks into (adverbial) free adjuncts in Modern English, while David Tizon-Couto's article looks at the development of left-dislocated strings. This finally brings us to elements with scope over clauses and larger units. Here, two papers deal with modality: Kousuke Kaita describes how the Old English preterite-present verb *dugan* developed into a modal auxiliary of ability in Middle English, while Christine Elsweiler examines the divergent use of the modal auxiliaries *will* and *shall* in Scots and Standard English. Finally, Reijirou Shibasaki's paper turns to an element with very wide – textual – scope, as it considers the newly emerging apo-koinou construction *that's (not) the bottom line is that*. We will now present each of the contributions in more detail.

In the first paper in this volume, **Bettelou Los** discusses diachronic developments involving the clausal constituent "subject". The paper presents a number of phenomena instantiating a change whereby the functionality of clause-initial adverbials as linkers to the immediately preceding discourse has been lost, and whereby

subjects have taken over this role of discourse linking. This increased the functional load of subjects and also led to more strategies to create subjects, such as cross-linguistically rare passives. One such strategy is the passive ECM-construction with verbs of thinking and declaring, as in *This mushroom is reported to have a lobster like flavor when cooked* <http://caldwell.ces.ncsu.edu/2015/01/336308>; another is the use of passives like *He was prescribed lithium*, where the subject fulfills the thematic role of recipient (see also Dreschler 2015). The paper then discusses two further strategies that facilitate so-called "permissive subjects": middles, as in *This book reads well*, and causative/ergative valency alternations of the type *Amazon shipped the order/The order has shipped*. Importantly, these permissive subjects go beyond providing an alternative expression for discourse linking adverbials. In other words, the phenomenon of "permissive subjects" seems to have acquired a dynamics of its own, far beyond the need to compensate for the loss of referential adverbials. Generally speaking, Los's paper brings home the point that syntactic change is not just a matter of changing form, but often also involves changing function.

Another paper within the domain of complementation is by **Nikolaos Lavidas**, who discusses (changes in) cognate noun constructions (CNCs) as a window on Bible translations. He looks at these CNCs (e.g. *to smile a disarming smile*) in Early Modern English, particularly in the first complete English translation of the Bible from the original Greek and Hebrew by William Tyndale. As this translation was produced during a period when there was a significant expansion of CNCs, it could be expected that the translation would follow the newest variant (with a zero-marked object), and not an earlier variant (with a cognate prepositional object). This expected development would be in line with the development observed by Rohdenburg (this volume) from objects marked by a preposition to zero-marked objects in the area of simple object structures. However, Lavidas notes that Tyndale's translation follows the early variant, thus deviating from the new tendency and from the source text. He points out that, in following this early/archaic variant, the CNCs behave according to the principles of a "sacral stamp", a continuation of an archaic tradition. In this case, the sacral stamp does not involve adhering to Christian and mainly Koine Greek linguistic features, but rather following earlier English rules. Lavidas furthermore mentions that Tyndale's tendency to use the earlier variant is in line with a general tendency in his work to use archaisms (as can be seen from his dispreference of auxiliaries). Accordingly, "CNCs function together with other archaic elements to establish an English biblical register rather than reflect a translation effect" (p. 72). Methodologically, this paper can be situated in the philological tradition, in that the author carefully examines a translator's output within a particular literary tradition. In that sense, the paper also shows that external factors such as text type can affect (in this case, have a conservative effect on) the development of particular syntactic structures.

Günter Rohdenburg's paper – situated at the interface between argument structure and verb complementation – studies two seemingly opposing trends in the history of English syntax (from the Early Modern English period to the present day). In the area of simple object structures comprising one zero-coded nominal complement, prepositional object structures, as in (7a) and (8a), have been replaced by direct object structures, as in (7b) and (8b). As a result, the functional load of the direct object has increased substantially (in a way similar to the increase in the functional load of the subject, as described by Los and Dreschler (2012) and Los (this volume)).

(7) a. He incited *to* a rebellion.
    b. He incited a rebellion.

(8) a. *…laugh not with him, … least thou gnash **with** thy teeth at the last…*
                                                      (*Early English Prose Fiction*, 1584)
    b. He gnashed his teeth.

In the area of more complex object structures, English has seen a number of reductive changes. Some complex structures have been phased out at the expense of less complex structures (e.g. *A wasp stung me on the finger > A wasp stung my finger*). Other complex structures such as the double object construction have seen a striking reduction, in that various classes of matrix verbs have given up the double object construction over the past few centuries (Rohdenburg 1995a: 109–112). With some classes of matrix verbs (verbs of separation such as *banish, bar* and verbs of dispossession such as *(be)reave, defraud*), the double object construction shifted to a construction consisting of object 1 + preposition + object 2; with other classes (directive verbs such as *command, counsel*), the construction was lost. This change has led to greater structural explicitness: the structure comprising a prepositional object is less opaque than the double object construction, and the narrower functional range of the double object construction (fewer matrix verbs) has effected a closer form–function fit. A third complex structure that has seen a striking reduction (a semantic specialization with several matrix verb classes either completely or partly lost) is complement clauses involving the frames S – V – O + *that*-clause and S – V – O + *to*-infinitive. Finally, what explains these differential evolutions? Rohdenburg argues that this differential development might be seen as a trade-off between the amount of processing required per syntactic frame and the degree of functional diversity or functional load. When amount of processing is limited (as in the direct object construction), the object will tend to increase its functional load (and take the place of structurally more complex objects (such as prepositional objects)). On the other hand, complex object constructions, which require quite a bit of processing, will reduce their functional load and be replaced by structures

that have greater structural explicitness. This trade-off is in line with Rohdenburg's Complexity Principle: in cases such a (7)–(8), the more explicit option (with the prepositional object) is not necessary because the simple object environment in not cognitively complex; conversely, more complex object environments (e.g. double object constructions) will prefer greater structural explicitness (e.g. object 2 replaced by a prepositional object).

Brian Lowrey examines the use of implicative causative constructions comprising the causative verbs *make, do*, and *cause* with two kinds of finite complements in Middle English: a simple finite clause [V+*that*] (9) and a more complex finite clause [V+NP+*that*]-clause (10):

(9)   *The clennesse and the fastynge of us freres / **Maketh** that crist accepteth oure preyeres.*
'The purity and fasting of us monks causes Crist to hear our prayers.'
<div align="right">(Chaucer, <i>Summoner's Tale</i>, 1883–1884)</div>

(10)   *For which, where as his people therbifore / Hadde loved him wel, the sclaundre of his diffame / **Made** hem that they hym hatede therfore.*
'For which reason, whereas his people had previously loved him well, the scandal of his bad reputation made them hate him for it.'
<div align="right">(Chaucer, <i>Clerk's Tale</i>, 729)</div>

The author addresses the question why these finite constructions are found in Middle English, in light of the fact that the replacement of finite *that*-clauses by *to*-infinitives with causative verbs would already have occurred in the transition from OE to ME. The question comes down to (i) why [V+*that*] was retained and (ii) why [V+NP+*that*] appears to be an innovation in Middle English. Searching through a sample of Middle English prose and poetic texts (totaling over 1,400,000 words), Lowrey suggests, first, that [V+*that*] in Middle English (in particular, *make+that*) is a kind of relic: it encodes "second causative" situations (with non-agentive causers and/or causees and stative complements) and occurs in much the same environments as in Old English. Although it will die out by Early Modern English, it is still marginally productive in Middle English in that it also starts occurring with the matrix verb *cause*, with the same secondary causative semantics. Second, in accounting for the occurrence of the [V+NP+*that*]-construction in Middle English (which was lost in Early Modern English), several factors are argued to be at work. In line with Traugott & Trousdale (2013: 54), Lowrey proposes (i) that the emergence of the construction (most often used with *make*) is a language-internal development most likely primed by semantically and syntactically related constructions, namely the [V+*that*] structure and [V+NP+Pred] structures (small clause constructions), and (ii) that the matrix verb *cause* is a new second causative in this construction, and

its presence is accounted for by analogy with *make*. On a more theoretical plane, Lowrey suggests that the new [V+NP+*that*]-construction in Middle English is a case of constructionalization, a new node being created within a constructional network of related constructions. In addition to the workings of analogy, Lowrey also examines whether an external factor, namely, contact influence with Norman French and Latin, may have been at play. This question still needs to be further explored, but any contact influence would have affected the [V+*that*]-construction rather than the [V+NP+*that*]-construction.

Following up on Lowrey's finding that finite complementation with causative *make* in Middle English shifts to infinitival complementation in Early Modern English, **Yoko Iyeiri** considers the variation between non-finite causative *to-* and bare infinitives in Early Modern English. After establishing that bare infinitives take up more than 80% of the relevant examples (in the active voice) by the second half of the seventeenth century, she focuses on the factors affecting the choice of bare infinitive and *to*-infinitive with causative *make*. An important factor in the choice between the two infinitival alternants is Rohdenburg's (1996: 151) Complexity Principle, which states that "In the case of more or less explicit grammatical options, the more explicit one(s) will tend to be favored in cognitively more complex environments" (see also Rohdenburg, this volume). Iyeiri's analysis, based on a four-million word corpus of Early Modern English Prose, shows that *to*-infinitives, which are indeed the more explicit grammatical option than bare infinitives, are preferred in "complex" environments, i.e. environments where the object of causative *make* consists of three words or more, as in (11), or where elements other than the object intervene between *make* and its complement, as in (12). Note, though, that towards the end of the Early Modern English period, bare infinitives come to be increasingly established with active causative *make* (arguably as a result of analogy with the more frequent pattern).

(11)   *This thinge maketh **me and many other** to meruayle.* (Sir Thomas Elyot, 1541)

(12)   *… which made **him shortly after** sing, Fortune my soe, &c.*

(Sir Robert Naunton, 1641)

Conversely, the bare infinitive is preferred when the object of causative *make* is a less complex personal pronoun rather than a more complex full noun phrase; this is, again, in keeping with the Complexity Principle. In structures involving coordinated infinitives, infinitival markings are likely to remain the same or be reduced in the second complement, as in (13):

(13)   *…which made him **to abide** in the Doctrine, and **follow** the Example of his Captain Crist Jesus …*   (Francis Patchet, 1678)

Iyeiri explains the loss of the infinitival marker *to* in the second complement in terms of its predictability, and it is thus in line with Givón's (1991: 87) Quantity Principle, whereby "Less predictable information will be given more coding material". She also suggests that the higher syntactic predictability of *to* in the second complement may be seen as a reduction of its cognitive complexity, which would, again, align the occurrence of the bare infinitive with the Complexity Principle. A final factor is the lexical item in the complement: verbs such as *appear* and *believe* have a preference for the bare infinitive, while *be* prefers marking by *to*; in Iyeiri's view, this distribution may be related to the stress pattern of the verb. All in all, Iyeiri shows that the variation between bare infinitive and *to*-infinitive is related to cognitive-functional principles. These factors ceased to be operative as bare infinitives became increasingly established with causative *make* in Modern English.

A third paper dealing with causative complementation structures is by **Mark Davies** and **Jong-Bok Kim**. Making use of several large corpora (*COHA, COCA, TIME*, and *GloWbE*), the authors examine the historical development of the "*into*-causative" construction [V+NP+*into*+V-*ing*] in American English from the 1800s. Examples are in (14).

(14)   a.   which was a venture that I never *was persuaded into undertaking* before
                                                        (*COHA*, FIC, 1869)
       b.   I have never yet been able to *reason* myself *into feeling* old.
                                                        (*COHA*, FIC, 1884)

They observe not only an overall increase in token frequency of the construction (from a normalized frequency in *COHA* of 4.13 pmw in the period 1840s–1860s to 14.99 in 1990s–2000s), but also an increase in lexical diversity (with up to 544 different matrix verbs that participate in the construction by the 2000s in *COHA* alone). The focus of the paper is on the changes in the semantic classes of the matrix verbs. Corpus analysis shows that matrix verbs with negative prosody (e.g. verbs of persuasion (*persuade, coax*); verbs of force (*force, drive*), and verbs of deception (*deceive*)) as well as so-called "neutral" verbs (*influence, lead, stimulate*) can be attested across the decades (note that, pace Rudanko 2005, the authors observe that these neutral verbs were already present 150–170 years ago, at nearly the same frequency as today). A new semantic class – and this appears to be a recent change – is that of "positive" matrix verbs such as *charm, enchant, romance*. As the authors point out, matrix verbs have from the beginning always included less negative verbs such as *persuade, coax, influence*, so "there was always the possibility of extending the construction to explicitly positive verbs such as *motivate, enchant*, and *love*" (p. 172). Another noteworthy extension is the spread of the construction from cases expressing direct causation, as in *Bill (X) talked Sue (Y) into paying (Z) for the meal*, to cases where the causation is much more indirect, as in (15):

(15)   It (X) has helped to *build* America (Y) *into exploring* new frontiers (Z).
<div align="right">(*COCA*, SPOK, 1994)</div>

The extension of the construction to positive matrix verbs (in direct causation) as well as the extension to matrix verbs occurring in indirect causation is in line with Goldberg's (1995, 2006) views on Construction Grammar: the verbs can occur in the "*into*-causative" construction as long they are compatible with the constructional meaning such that X exerts some force or coercion on Y, resulting in Z. As in Lowrey's paper, analogy as a language-internal factor of change is seen to play a role here as well. In a final note, Davies and Kim point to the importance of robust corpus-based data. Without the availability of (large) corpora, this type of research would have been impossible, or would have yielded less reliable results.

Continuing on the theme of non-finite clauses, **Carla Bouzada-Jabois**'s paper addresses the development of free adjuncts in Late Modern English, making use of corpus data from the *Penn Parsed Corpus of Modern British English* and of quantitative analyses of Early Modern English (Rio-Rey 2002) and Present-day English (Kortmann 1991). Free adjuncts are defined as "typically nonfinite constructions conveying adverbial meaning with respect to the main clause", as in (16):

(16)   They were going away, and we were going after them, *firing at them too.*
<div align="right">(*PPCMBE1*, HOLMES-TRIAL-1749,69.1289)</div>

Against the background of a slight overall decline of free adjuncts from Early Modern English to Present-day English (from a normalized frequency of 33.6 per 10,000 words to 29.5), the author provides a quantitative description of (changes in) the features of free adjuncts. With regard to the *head-element*, free adjuncts show an increasingly higher share of present participles (at the expense of perfect and past participles). As to *position* in the sentence, free adjuncts can take up initial, medial, and final position. Free adjuncts show an increasing preference for sentence-final position; this increase correlates with a decrease of medial free adjuncts (as the share of initial free adjuncts remains relatively constant over time). *Augmentation* (by conjunctions), then, is seen to increase proportionally in the Late Modern English period, but decreases between Late Modern English and Present-day English. As regards the *semantic typology* of free adjuncts, Bouzada-Jabois adopts Kortmann's (1991: 121–133) classification in terms of "least informative" and "most informative" semantic relations (e.g. simultaneity and concession, respectively). Her findings show that in Late Modern English, the relative frequencies of more or less informative semantic relations are very similar, and that it is only in Present-day English that free adjuncts conveying more informative relations become the preferred type. Semantic type is also seen to correlate with position: the least informative relations favor end position, while the most informative relations favor initial

position. Finally, with regard to *text type*, Bouzada Jabois' results show a preference for free adjuncts in writing-related text types.

Another phenomenon within the clausal periphery is examined in **David Tizon-Couto**'s paper on the diachronic development of Left Dislocated constructions, i.e. syntactic strings that include a left-dislocated constituent and a co-referring resumptive element in the subsequent clause. A Present-day English example is (17), where *this girl* is the left-dislocated constituent and *she* is the resumptive element.

(17)   *This girl* this morning *she* threw a wobbly.   (Biber et al. 1999: 956)

The data were extracted from the *Penn-Helsinki Parsed corpora of Middle English* (*PPCME2*), *Early Modern English* (*PPCEME*), *Modern British English* (*PPCMBE*), and the *Parsed Corpus of Early English Correspondence* (*PCEEC*). Against the background of an overall decrease of Left Dislocated strings from Middle English into Modern English (with biblical texts being the only genre that do not show a decline), the paper identifies the factors that determine the structural complexity of the Left Dislocated constituent (measured in terms of its length). Relevant for research into syntactic change is the factor or independent variable "Period", which the author reports to have a significant effect on length: there is a tendency for the complexity of the Left Dislocated constituent to decrease in time, except for the last period of Late Modern English.[7] The author also explores the link between the general decrease of Left Dislocated NPs and genre (and in particular speech-related texts; see Culpeper & Kytö 2010). It appears that the decrease across the Modern English period is only significant for "writing-related" texts and, interestingly, also for "speech-like" texts (e.g. letters and diaries). The latter finding is somewhat surprising in that one might expect Left Dislocated strings to become markers of unintended orality. Rather, it is the genres where writers deliberately attempt to reproduce orality that fail to show a significant decrease of Left Dislocated string frequency. In other words, it does not seem likely that Left Dislocated strings constitute traces of unintended orality. In general, this paper, again, shows the effect of text/genre on the diachronic development of syntactic structures.

Then follow two papers on modality. **Kousuke Kaita**'s paper examines how the Old English preterite-present verb *dugan* 'to avail' developed into a modal auxiliary of ability ('can') in Middle English. The modal was lost before the Modern English period (though, according to Nagle & Sanders 1988: 253, it survives "as a modal in some regional varieties of British English"). Data have been extracted from a series of Old English and Middle English prose and verse texts (including several medical texts). Kaita's main hypothesis is that the auxiliation of *dugan* can be accounted

---

7.   The author admits, however, that more data are needed to confirm this diachronic trend.

for by similarity with Old English *magan* at the morphological, semantic, syntactic, and textual level. Aside from the fact that, morphologically, the similarity between both verbs could already be observed at the PIE stage, OE *dugan*/ME *douen* shares features with *magan/mouen* as to infinitival and present-participle forms. Semantically, *dugan* describes some positive potential for its animate subject of being 'good', 'useful', a meaning similar to *magan*'s 'to be strong, efficacious'. The majority of *dugan* examples, with positive semantics, occur in medical texts, where they typically take the *to*-infinitive and the preposition *wið* 'against' (see (18)); these are semantic and syntactic features *dugan* shares with lexical *magan* in medical texts.

(18) *Wiþ   latre meltunge, olisatrum   hatte wyrt*, *seo*   **deah**   *to*
      against late   digestion olustratum called wort   which is.good to
      drincanne.   (*Lch* II (2) 34.2.1, s. x med.)
      drink
      'For late digestion; a wort hight olustratum, which is good to drink.'
                                                             (Cockayne 1865: 239)

These shared characteristics of *dugan* and *magan* were maintained in Middle English medical texts, which contributed further to *dugan*'s auxiliation. The auxiliation proper involved a reanalysis from Middle English *douen* + *to*-infinitive in impersonal use + dative object (19) to a personal construction with nominative subject (20). Semantically, *douen*'s notion of 'propriety' (developed from 'positive evaluation') extended to 'ability'.

(19) *as meiden* **deh**   *to beonne*
      as maiden is.proper to be
      'as it is proper for a maiden to be'
                                      (*St. Juliana* 487 (MS Bodley 34), c1225(?c1200))[8]

(20) *Fight he aght   ai   quils   he* **dught**,
      fight he ought always while   he was.able.to
      And fle   quen he langer   ne   **moght**.
      and flee when he no   longer could
      'He ought to have always fought while he was able to, and fled when he could no longer (fight).' (*Cursor Mundi* 23771–23772 (MS Cotton), a1400(a1325))

This paper neatly shows the workings of such language-internal factors as analogy and reanalysis, as well as the importance of textual genre, in effecting (morpho) syntactic change.

---

8. The dating of the Middle English examples cited is based on that provided in *Middle English Dictionary (MED)*.

Christine Elsweiler investigates the historical development of *will* and *shall* in Older Scots, with a focus on the divergent use of the modal auxiliaries *will* and *shall* in Scots and Standard English. In this qualitative and quantitative study, she draws on corpus data from the *Helsinki Corpus of Older* Scots (*HCOS*). In particular, the following research questions form the basis of her study:

1. Do the *HCOS* data show a decline of first-person volitional *will*?
2. Can a concomitant rise of first-person volitional *shall* be observed?
3. Is there a rise in predictive uses of first-person *will*?
4. Is there evidence for the use of *will* in the second and third person to express obligation?

Addressing these questions, Elsweiler's shows that across the Older Scots period, there is a significant decrease in volitional *will*, not only in the first person, but also in the third person. A concomitant rise of first-person volitional *shall* is, however, not observed. With regard to first-person predictive *will*, no increase can be observed; the overall increase in predictive *will* occurs in the third person; there are no marked changes in predictive *shall* generally. As to the final research question, several instances of second- and third-person *will* were found in the indeterminate category 'prediction/obligation'. In addition to tracing changing frequencies of the different uses of Scots *will* and *shall*, she also provides information about the typical genres in which the uses occur. Elsweiler's data analysis furthermore reveals that (i) the general drop in frequency of volitional *will* and the rise of predictive *will* in Older Scots match the development of the modal meanings of *will* in Early Modern English, and (ii) that for *shall*, the obligation meaning is predominant throughout the sub-periods, in spite of a general decline of this use.

The last contribution addresses the marker with the broadest scope. **Reijirou Shibasaki** examines the (recent) diachrony of the apo-koinou construction *that's (not) the bottom line is that*. Data are extracted from the *Corpus of Historical American English* (*COHA*) and from the *Corpus of Contemporary American English* (*COCA*). It is argued that this construction has emerged from the sequential use of the anaphoric construction *that's (not) the bottom line*, allowing the speaker to summarize his/her preceding statement(s), and the cataphoric construction *the bottom line is (that)*, allowing the speaker to introduce or anticipate upcoming talk. Shibasaki not only focuses on constructions comprising the item *the bottom line*, but more generally, considers the sequential use and apo-koinou constructions involving selected shell nouns (*point, issue, problem, question, truth, thing*). While sequential uses often contain *not*, apo-koinou constructions only emerge when *not* is not used. Furthermore, apo-koinou constructions do not refer back to previously mentioned utterances but refer forward to subsequent utterances. A final section of the paper describes the emergence of the apo-koinou construction as a case

of constructionalization (see Traugott & Trousdale 2013): it is formally different from its component constructions, and it converges the semantics of its component constructions into cataphoric reference. It is also argued that the formation of more recent apo-koinou constructions, such as the one involving *bottom line*, is facilitated by its analogy with similar, more frequent constructions (involving the shell nouns *thing*, *problem*). Once constructions at lower levels of schematicity such as *that's the thing is (that)* and *that's the problem is (that)* have given rise to a more schematic construction "that's the X is (that)", this more schematic construction, in turn, may sanction newer construction such as *that's the bottom line is (that)*. More generally, this paper is another example of the importance of considering performance-related data as well as the working of the factors analogy and frequency in describing syntactic change.

In sum, then, the papers in this volume cover a wide range of interrelated syntactic phenomena, from the history of core arguments, to complements and non-finite clauses, elements in the clause periphery, as well as elements with potential scope over complete sentences and even larger discourse chunks. In addition, the papers testify to an increasing awareness that even some of the most central phenomena of syntax – and the way they develop over time – are best understood by taking into account their communicative functions and the way they are processed and represented by speakers' cognitive apparatus. In this light, it is fair to say that historical syntax, and historical linguistics in general, is witnessing a convergence between formerly distinct linguistic frameworks and traditions. Much of the descriptive apparatus in this volume is clearly of generative stock and the descriptive finesse currently expected also of historical research is in large measure thanks to the rigorous theoretical work by formal syntacticians. At the same time, historical syntax is not practiced in its purest form, which may – in the end – prove to have been only a useful idealization, but is increasingly being complemented with the insights of functional-cognitive approaches. With this fusion of traditions, we believe the trend is undeniably towards a richer and more broadly informed understanding of syntactic change and the history of English.

Acknowledgements

References

Andersen, Henning. 2001. Actualization and the (uni)directionality of change. In *Actualization: Linguistic Change in Progress* [Current Issues in Linguistic Theory 219], Henning Andersen (ed.), 225–248. Amsterdam: John Benjamins. doi:10.1075/cilt.219.11and

Bergs, Alexander & Brinton, Laurel (eds). 2012. *English Historical Linguistics: An International Handbook*, 2 Vols. Berlin: De Gruyter Mouton.

Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Bowern, Claire & Evans, Bethwyn (eds). 2015. *The Routledge Handbook of Historical Linguistics*. London: Routledge.

Brinton, Laurel (ed.). 2017. *English Historical Linguistics: Approaches and Perspectives*. Cambridge: CUP.

Brinton, Laurel & Bergs, Alexander (eds). 2017. *The History of English: Historical Outlines from Sound to Text*. Berlin: De Gruyter Mouton.

Brems, Lieselotte. 2010. Size noun constructions as collocationally constrained constructions: Lexical and grammaticalized uses. *English Language and Linguistics*. 14: 83–109.

Bybee, Joan. 2003. Mechanisms of change in grammaticalization: The role of frequency. In *The Handbook of Historical Linguistics*, Brian Joseph & Richard Janda (eds), 602–623. Oxford: Blackwell. doi:10.1002/9780470756393.ch19

Bybee, Joan. 2007. *Frequency of Use and the Organization of Language*. Oxford: OUP. doi:10.1093/acprof:oso/9780195301571.001.0001

Canale, Michael. 1978. Word Order Change in Old English. PhD dissertation, University of Toronto.

Culpeper, Jonathan & Kytö, Merja. 2010. *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: CUP.

De Smet, Hendrik. 2009. Analysing reanalysis. *Lingua* 119: 1728–1755. doi:10.1016/j.lingua.2009.03.001

De Smet, Hendrik. 2013. *Spreading Patterns: Diffusional Change in the English System of Complementation*. Oxford: OUP.

D'hoedt, Frauke & Cuyckens, Hubert. 2017. The development of the *as*-Secondary Predicate Construction: Constructionalization and internalization. *Language Sciences* 59: 16–35.

Dreschler, Gea. 2015. *Passives and the Loss of Verb-second: A Study of Syntactic and Information-structural Factors*. Utrecht: LOT.

Fischer, Olga. 1988. The rise of the *for NP to V* construction: An explanation. In *An Historic Tongue: Studies in English Linguistics in Memory of Barbara Strang*, Graham Nixon & John Honey (eds), 67–88. London: Routledge.

Fischer, Olga. 2007. *Morphosyntactic Change: Functional and Formal Perspectives*. Oxford: OUP.

Fischer, Olga. 2013. The role of contact in English syntactic change in the Old and Middle English periods. In *English as a Contact Language*, Daniel Schreider & Marianne Hundt (eds), 18–44. Cambridge: CUP.

Fischer, Olga & van der Wurff, Wim. 2006. Syntax. In *A History of the English Language*, Richard Hogg & David Denison (eds), 109–198. Cambridge: CUP. doi:10.1017/CBO9780511791154.004

Fischer, Olga, De Smet, Hendrik & van der Wurff, Wim. 2017. *A Brief History of English Syntax*. Cambridge: CUP.

Foster, Tony & van der Wurff, Wim. 1995. The survival of object–verb order in Middle English: Some data. *Neophilologus* 79: 309–327   doi:10.1007/BF00999787

Givón, Talmy. 1991. Isomorphism in the grammatical code: Cognitive and biological consideration. *Studies in Language* 15: 85–114.   doi:10.1075/sl.15.1.04giv

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago IL: University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at Work: Constructionist Approaches in Context*. Oxford: OUP.

Harris, Alice & Campbell, Lyle. 1995. *Historical Syntax in Cross-linguistic Perspective*. Cambridge: CUP.   doi:10.1017/CBO9780511620553

Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: CUP.   doi:10.1017/CBO9781139004206

Hilpert, Martin & Gries, Stefan T. 2016. Quantitative approaches to diachronic corpus linguistics. In Kytö & Pahta (eds), 36–53.

Hogg, Richard (ed.). 1992–2001. *The Cambridge History of the English Language*, 6 Vols. Cambridge: CUP.   doi:10.1017/CHOL9780521264747

Hopper, Paul & Traugott, Elizabeth. 2003. *Grammaticalization*. Cambridge: CUP.

Joseph, Brian & Janda, Richard (eds). 2003. *The Handbook of Historical Linguistics*. Oxford: Blackwell.   doi:10.1002/9780470756393

Kortmann, Bernd. 1991. *Free Adjuncts and Absolutes in English: Problems of Control and Interpretation*. London: Routledge.

Kytö, Merja & Pahta, Päivi (eds). 2016. *The Cambridge Handbook of Historical Linguistics*. Cambridge: CUP.

Langacker, Ronald. 1977. Syntactic reanalysis. In *Mechanisms of Syntactic Change*, Charles Li (ed.), 57–139. Austin TX: University of Texas Press.

Lightfoot, David. 1979. *Principles of Diachronic Syntax*. Cambridge: CUP.

Lightfoot, David. 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge MA: The MIT Press.

Lightfoot, David. 1999. *The Development of Language: Acquisition, Change, and Evolution*. Oxford: Blackwell.

López-Couso, María José. 2016. Corpora and online resources in English historical linguistics. In Kytö & Pahta (eds), 127–145.

Los, Bettelou. 2012. The loss of verb-second and the switch from bounded to unbounded systems. In *Information Structure and Syntactic Change in the History of English*, Anneli Meurman-Solin, María José López-Couso & Bettelou Los (eds), 21–46. New York NY: OUP.   doi:10.1093/acprof:oso/9780199860210.003.0002

Los, Bettelou. 2015. *A Historical Syntax of English*. Edinburgh: EUP.

Los, Bettelou & Dreschler, Gea. 2012. The loss of local anchoring: From adverbial local anchors to permissive subjects. In Nevalainen & Traugott (eds), 884–898.

Mair, Christian. 2016. Audio recordings. In Kytö & Pahta (eds), 146–163.

Mugglestone, Linda (ed.). 2006. *The Oxford History of English*. Oxford: OUP.

Nagle, Stephen J. & Sanders, Sara L. 1998. Downsizing the preterite-presents in Middle English. In *Advances in English Historical Linguistics*, Jacek Fisiak & Marcin Krygier (eds), 253–261. Berlin: Mouton de Gruyter.

Nevailainen, Terttu & Raumolin-Brunberg, Helena. 2003. *Historical Sociolinguistics*. London: Pearson.

Nevalainen, Terttu & Traugott, Elizabeth (eds). 2012. *The Oxford Handbook of the History of English*. Oxford: OUP. doi:10.1093/oxfordhb/9780199922765.001.0001

Pintzuk, Susan & Taylor, Ann. 2011. The independence of information status effects and the syntactic change from OV to VO in the history of English and Icelandic. Paper presented at the 13th Diachrony in Generative Syntax Conference (DiGS), University of Pennsylvania, 2–5 June.

Río-Rey, Carmen. 2002. Subject control and coreference in Early Modern English free adjuncts and absolutes. *English Language and Linguistics* 6(2): 309–323. doi:10.1017/S1360674302000254

Rohdenburg, Günter. 1995a. Betrachtungen zum Auf- und Abstieg einiger präpositionaler Konstruktionen im Englischen. *NOWELE* 26: 67–124. doi:10.1075/nowele.26.05roh

Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7: 149–182. doi:10.1515/cogl.1996.7.2.149

Rudanko, Juhani. 2005. Lexico-grammatical innovation in current British and American English: A case study on the transitive *into -ing* pattern with evidence from the Bank of English Corpus. *Studia Neophilologica* 77: 171–187. doi:10.1080/00393270500384791

Simpson, Jane. 2002. From common ground to syntactic construction: Associated path in Warlpiri. In *Ethnosyntax: Explorations in Grammar and Culture*, Nick Enfield (ed.), 287–307. Cambridge: CUP.

Stein, Dieter. 1990. *The Semantics of Syntactic Change: Aspects of the Evolution of 'do' in English*. Berlin: Mouton de Gruyter. doi:10.1515/9783110846829

Traugott, Elizabeth & Trousdale, Graeme. 2013. *Constructionalization and Constructional Changes*. Oxford: OUP. doi:10.1093/acprof:oso/9780199679898.001.0001

Trousdale, Graeme. 2017. Syntax. In Brinton & Bergs (eds.), 102–122.

van de Pol, Nikki & Cuyckens, Hubert. 2014. The diffusion of English absolutes. In *Corpus Interrogation and Grammatical Patterns* [Studies in Corpus Linguistics 63], Kristin Davidse, Caroline Gentens, Lobke Ghesquière & Lieven Vandelanotte (eds), 265–294. Amsterdam: John Benjamins.

van de Pol, Nikki. 2016. The Development of the Absolute Construction in English: Between Bird's Eye View and Magnifying Glass. PhD dissertation, KU Leuven.

van Gelderen, Elly. 2016. Generative frameworks and approaches. In Kytö & Pahta (eds), 79–94.

van Kemenade, Ans. 1987. *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.

van Kemenade, Ans. 2009. Discourse relations and word order change. In *Information Structure and Language Change*, Roland Hinterhölzl & Svetlana Petrova (eds.), 191–220. Berlin: Mouton de Gruyter.

van Kemenade, Ans & Los, Bettelou (eds). 2006. *The Handbook of the History of English*. Chichester: Wiley-Blackwell. doi:10.1002/9780470757048

Viana, Vander, Sonia Zyngier & Geoff Barnbrook (eds). 2011. *Perspectives on Corpus Linguistics* [Studies in Corpus Linguistics 48]. Amsterdam: John Benjamins. doi:10.1075/scl.48

"Permissive" subjects and the decline of adverbial linking in the history of English

Bettelou Los
University of Edinburgh

Earlier work by Los and Dreschler (2012) and Komen et al. (2014) offered quantitative evidence of a decline in clause-initial adverbials as discourse linkers in the history of English, and argued that subjects have taken over much of the function of discourse linking that was earlier performed by adverbials. The greater functional load of the subject called for more flexibility in which types of thematic roles could be expressed by subjects, and for more strategies to create subjects, such as crosslinguistically rare passives. The present paper draws attention to another mechanism that facilitates "permissive subjects" in Present-day English: causative/ergative valency alternations of the type *Amazon shipped the order/The order has shipped.* I present the morphosyntactic origin of the alternation and report in more detail on the workings of discourse linking in Old English texts, explaining why "permissive subjects" were not required at that stage.

1. Introduction

There is a growing body of research indicating that the functional load of the subject in English has increased from Middle English onwards. In German narratives, subjects primarily encode protagonists, mind-possessing agents who play a role in the events that are the narrative's concern. There is evidence from psycholinguistic experimental work that English speakers differ from German speakers in that they are more inclined to mention non-protagonists, like *The wind* in (1):

(1)   A young man is surfing. *The wind* is blowing him off the board.

(Carroll et al. 2004: 190)

Although the German equivalent of the verb *blow* allows a non-animate subject like *the wind*, German speakers in this experiment tend to keep focused on the protagonist, the young man, with descriptions equivalent to Present-day English (PDE) *He is blown off the board* (ibid.). Carroll et al. (2004) ascribe this finding to a difference in narrative perspective: the German speakers tell a story from the perspective of the protagonist, while the English speakers take an outsider's view, like a camera. Los (2012) hypothesizes that this difference may have come about relatively recently, with Old English (OE) being more like German and Dutch. If PDE subjects are more likely to encode non-protagonists than in earlier times, this would increase the functional load of subjects.

A second reason why the functional load of subjects can be hypothesized to have increased in the course of the history of English is another contrast between PDE and German: the finding by Rohdenburg (1974), and after him by Hawkins (1986), that PDE allows subjects that would be expressed by adverbials in German, and, we may add, in Present-day Dutch (see (2b)):

(2) a.   This *loses* us the best centre forward.

                    (Hawkins 1986: 58–61, from Rohdenburg 1974: 11)

    b.   *Hiermee zijn we onze beste spits   kwijt.*
        herewith are we our   best   centre-forward lost
        'With this, we have lost our best center forward.'

Los and Dreschler (2012) and Dreschler (2015) have shown that adverbial links of the type *Hiermee* 'with this' in (2b) have decreased since Middle English.

Komen et al. (2014) present quantitative support for the idea that the functional load of subjects has increased: PDE narratives have higher rates of inanimate subjects than Old English (OE), which is what we would expect if the subject starts to express non-protagonists like *The wind*, as in (1), and discourse links like *This*, as in (2). Further support comes from the higher rates of subject-switch between clauses.

High rates of subject-switch can be expected to require more topic-reactivation by full nominals (names and descriptions like *Her husband* rather than pronominal subjects like *He*), as the subject is no longer reserved for protagonists only, which leads to a lower ratio of pronominal versus nominal subjects. For the same reason, we would expect higher rates of subject ellipsis in OE (as has also been found for the contrastive English/German study of Carroll & Lambert 2003): if the subject's primary function is to express the protagonist, it should be easier to recover the referents of elided subjects than if the subject is routinely used for other functions. These hypotheses were confirmed by a pilot study in Los (2009), as well as a larger scale investigation in Komen et al. (2014).

The examples in Rohdenburg (1974) and Hawkins (1986) show that subjects in English are extremely "permissive" compared to those in related languages such as

German, in that verbs normally associated with animate subjects, like *lose* in (2), show up with inanimate subjects in a variety of semantic roles. The focus of the present paper is on how the increase in functional load of the subject is accompanied, and accommodated, by changes elsewhere in the grammar.

The paper is structured as follows. Section 2 discusses "permissive subjects" in more detail as a specialty of PDE. Section 3 shows the various syntactic positions available in OE and how they could be exploited to distribute given and new information over the clause; this section highlights the discourse-linking role of first-position adverbials, usually encoded as prepositional phrases (PPs). Section 4 presents the phenomenon of "permissive subjects" in the context of flexible argument structure, and explores some of the accidental losses in the morphology that led to flexible argument structures becoming a productive process. Over time, given enough examples, such patterns apparently attracted so many members that the possibilities for verbs developing alternative argument structures snowballed and acquired a dynamics of their own. As a result, the valency alternation no longer needed to be sanctioned by usage for each individual verb; instead, groups of verbs came to participate in the alternation in their entirety, so that in PDE any verb in *-ify* or *-en* has acquired both intransitive and transitive uses. Sections 5 and 6 provide some further discussion about the difficulty of investigating the increase of the functional load of the subject in terms of a straightforward scenario of replacement or competition: adverbial discourse links have themselves acquired a different role in that they are now part of "late subject" constructions, which are mainly presentational (Section 5). PDE is also very different from its West Germanic cousins in that it does not require tail–head linking in narratives, which may have meant that the decline of adverbial discourse links triggered the loss of tail–head linking and hence decreased the need for discourse linking itself. Section 7 concludes.

2.   "Permissive subjects"

Exactly how cross-linguistically unusual English subjects are compared to other European languages is brought home by work on machine translation. Compare these English/Portuguese pairs, where English can have the same verb for ergative as well as causative clauses, but Portuguese has to go for two different verbs entirely (all examples from Santos 1988); the Portuguese verbs are in bold:

(3)   a.   he returned – *ele **voltou***
       b.   he returned the book – *ele **devolveu** o livro*

(4)   a.   he quit – *ele **desistu***
       b.   he quit the job – *ele **deiou** o trabalho*

The same goes for actives and passives in such ergative/causative pairs:

(5)   a.   the program runs – *o programa **funciona***
      b.   the program is run – *o programa **é executado***

These examples show that the "permissive subjects" in PDE are accommodated by a flexible argument structure (of verbs like *return, quit*, and *run*); we will return to such valency phenomena in Section 4.

Another source of problems is Exceptional Case-Marking (ECM) constructions like *I intend the program to work with two different kinds of text files, I have shown the problems to be caused by X*, where the direct objects *the program* and *the problems* do not receive their theta-role from the verb (cf. **I intend the program*, *#I have shown the problem*) but from the lower verb, the infinitive. A small number of verbs (e.g. *intend, want, desire*) allow both ordinary subject-controlled infinitives, where the controller is an AGENT, as in (6a) and (7a), and ECM, in which, when passivized, non-agentive direct objects will become subjects, as in (6b) and (7b). ECM constructions with verbs of thinking and declaring are cross-linguistically rare, and the flexible valency of these verbs – subject-control and ECM – often cannot be duplicated in languages other than English. Verbs of thinking and declaring will tend to allow the subject-control construction but such constructions do not have a direct object and hence will not allow passivization; the same holds for English subject-control constructions: *I intend to travel light* versus **To travel light/*Traveling light is intended by me*). As the ECM-construction does contain a direct object, these same verbs can be found passivized in English; but in many other languages, a different verb will have to be selected to translate such English passive ECM constructions:

(6)   a.   I intend – *eu **tenciono***
      b.   the program is intended – *o programa **destina-se***
           cf. The program is intended to work with two different kinds of text files.

(7)   a.   the problem shows – *o problema **aparece***
      b.   the problem was shown – *o problema **foi mostrado***
           cf. The problem was shown to be in the local loop.

Like the proliferation of valency alternations, which will be dealt with in Section 4 below, the ECM-construction with verbs of thinking and declaring is recent, dating from late Middle English (see, e.g., Warner 1982; Fischer 1989; Fanego 1992; Los 2005; Dreschler 2015). All authors note that, at the emergence of the construction with these verbs, there is an asymmetry between actives and passives that persists into the present day, with many verbs not having an acceptable active counterpart. A PDE example is (8); see also Postal (1974).

(8)   [The picture shows a fruiting structure of the Hericium erinaceus fungus. Commonly called a lion's mane,] this mushroom is reported to have a lobster like flavor when cooked.   http://caldwell.ces.ncsu.edu/2015/01/336308/ (accessed on 19 October, 2015, via WebCorp, <http://www.webcorp.org.uk>)

Example (8) is totally acceptable, but the ECM-construction that is its active counterpart, *People reported this mushroom to have a lobster like flavor*, is not. Birner and Ward (2002) label the passive ECM-construction an "information-packaging" device, its primary function being to allow discourse-old information to be expressed as a subject. PDE examples like (8) demonstrate that passive ECM-constructions with these verbs assign subject status to inanimate, non-agentive entities.

I have claimed in earlier work (Los 2005: 257; Los 2009) that an entity like the *mushroom* in (8) would tend to be expressed by an adverbial in Dutch and German. As these languages did not develop the ECM with verbs of thinking and declaring, they will have to resort to constructions equivalent to *About this mushroom one reports / it is reported* [lit. *reports one / is it reported*] *that it has a lobster like flavor when cooked.* As the rate of first-position adverbials expressing discourse links declines from Middle English onwards, passive ECMs are a way around the restrictions on how such "unmarked themes" (sentence beginnings conveying "given" information without any additional sense of prominence or contrast; see, e.g., Downing & Locke 1995, following Halliday 1967, 1968) could be expressed – by subjects.

The next section will look at how King Alfred dealt with the ECM-constructions he found in the Latin *Vorlage* of his translation of Boethius' *De Consolatione Philosophiae*, the type of argumentative text that is the natural habitat for verbs of thinking and declaring, and how Ælfric uses one such verb, *tocnawan* 'understand, interpret', in his original OE argumentative prose. The flexibility of OE syntax fitted the purposes of information structure like a glove, particularly when it came to making links with the immediately preceding discourse; when this flexibility was lost, the default expression for such links was the subject.

3.   Adverbial positions and discourse linking in Old English

When King Alfred translated Boethius' *De Consolatione Philosophiae* into Old English, he came across ECM-constructions with passives of verbs of thinking and declaring, as in (9a), (10a), and (11a), where modern translations, as in (9b), (10b), and (11b), routinely use a passive ECM; the verb and its dependent infinitive are given in bold in the a-sentences and in italics in the b-sentences. King Alfred's translation makes no attempt to create a non-finite construction but has active verbs of thinking and declaring (*cweðað* 'say' in (9c), *wenað* 'think' in (10c), *ðincð*

'seem' in (11c)) followed by a finite (subjunctive) clause. These OE translations accordingly have to add animate subjects that are not in the *Vorlage*, all with generic reference: *we* in (9c), *hi* (= *men* '*people*') in (10c). But even if the subject of the ECM is animate, as in (11), Alfred has to resort to a finite construction, as the non-finite construction is not available to him.

(9)   a.   *Uti summa cardo atque causa expetendorum omnium bonitas **esse iure** **credatur**.*   (<http://data.perseus.org/citations/urn:cts:latinLit:stoa0058.
stoa001.perseus-lat1:3.P10>; O'Donnell 1990)

   b.   Goodness *is* rightly *believed to be* the sum and hinge and cause of all things desirable.   (James 1897)

   c.   *Forðæm   we   cweðað   þæt   þæt   hehste   good   sie   se*
therefore   we   say   that   that   highest   good   is.SUBJ   the
hehsta   hrof   eallra   gooda   and   sio   hior   ðe   eall   good   on
highest   roof   of-all   good   and   the   hinge   that   all   good   on
hwearfað,   and   eac   ðæt   þing   þe   mon   eall   good   fore   deð
turns   and   also   that   thing   that   one   all   good   for   does.
<Bo 34.88.10>[1]

(10)   a.   *Idcirco enim sufficientia petitur, quoniam bonum **esse iudicatur**.*
(<http://data.perseus.org/citations/urn:cts:latinLit:stoa0058.
stoa001.perseus-lat1:3.P10>; O'Donnell 1990)

   b.   For the very reason why independence is sought is that it *is judged good*.
(James 1897)

   c.   *þi   men   secaþ   god   genog   þe   hi   wenað   þæt*
therefore   people   seek   good   sufficiency   that   they   think   that
ðæt   sie   þæt   hehste   good.   <Bo 34.88.1>
that   is.SUBJ   that   highest   good

(11)   a.   *cui sententiae consequens est ut tum demum grauioribus suppliciis urguean-
tur cum impuniti **esse creduntur***   (<http://data.perseus.org/citations/
urn:cts:latinLit:stoa0058.stoa001.perseus-lat1:4.P4>; O'Donnell 1990)

   b.   from which point of view it follows that the wicked are afflicted with more severe penalties just when they *are supposed to escape* punishment
(James 1897)

---

1.   The reference to an OE text enclosed in < > follows the system of short titles as employed in Healey & Venezky (1985 [1980]) (in turn based on the system of Mitchell et al. 1975, 1979). It is identical to the TEI reference in the *DOEC* (= *Dictionary of Old English Corpus*, also known as the *Toronto Corpus*), which means that line numbers refer to the beginning of the sentence rather than the line in which the relevant structure occurs.

    c.   *Git hit gebyreð þæt ðe   ðincð ðæt þa orsorgan   bioð*
        yet it is.right that to.you seems that the unpunished are
        ungesæligran þonne þonne þa gewitnodan.   <Bo 28.121.2>
        unhappier   than   than the punished

Of particular interest to me are cases where verbs of thinking or declaring are used to express an opinion about a topic that is newly introduced in the immediately preceding discourse, like *this mushroom* in (8). This seems to be the main purpose of passive ECMs; note how the passive ECM in the Modern English translation in (9b) also starts with the previously introduced concept of *bonitas* as its topic and goes on to add a comment about it: *Goodness is rightly believed to be the sum … of all things desirable.* Alfred finds himself unable to use *bonitas* as the subject of a main clause which has as its main verb a verb of thinking and declaring, but instead embeds it as the subject of the subclause. He could have rephrased *goodness* as an adverbial (*About goodness we say…*), which would have fitted the function of the first position of the main clause in a V2 language well, as a position for expressions referring back to material introduced in the immediately preceding discourse, but that is not what he does. He fills this first position of the main clause with the backward-linking *forðæm* 'for that (reason)' instead. The effect is that the main clause is low on informational content – all the information is in the subclause.

    To see what happens in original rather than translated OE prose, the remainder of this section looks at one particular verb of thinking, *tocnawan* 'know, acknowledge, understand, interpret', a verb very typical of the vocabulary used by Ælfric in his homilies (argumentative, exegetical, discourse). All the examples below are from his works. The information about word order for these examples employs the traditional labels prefield, middle field, and postfield to refer to various salient domains within the architecture of the West Germanic clause.

    Using this terminology carries the risk that it imposes standard structural notions from what we know about Present-day Dutch and German onto OE syntax, so we need a number of disclaimers here: (i) these labels abstract away from the reanalysis to VO base word order which is already in evidence in Old English (see, e.g., Pintzuk & Taylor 2006; Taylor & Pintzuk 2012), but this paper is mostly concerned with the left edge of the main clause rather than the underlying position of the object; and (ii), although the current consensus is that there is finite verb movement in OE (following seminal work by van Kemenade 1987), we need to stress that this movement is not as syntactically motivated as V2 in Dutch and German where its main function is to signal clause type (main clauses, conditional clauses). The consensus is that there are two landing sites for verb movement in OE, as identified by, e.g., Haeberli (2002) and van Kemenade & Westergaard (2012), rather than the single position usually assumed for Dutch and German. Our concern is the

lower of these positions, where the finite verb may end up, on the surface, as V3, particularly if the first constituent is an object or an adverbial and the subject is a pronoun (see again van Kemenade 1987), as in (13) below, where we will discuss this matter further.

The argument structure of *tocnawan* calls for a subject and an object, and Ælfric's particular usage also calls for an adverbial: a person (= subject) understands something (= object) about something (= adverbial) or because of something (= adverbial). The first example, (12), is of *tocnawan* in a subclause, which shows, in this particular case, the basic SOV order of Subject – Object – Verb; the pronominal adverb *þæron* 'about that' is probably a postmodification of *sum andgit* inside the object-NP and does not occupy an adverbial slot at the level of the clause. It is cataphoric, linking to the following *how*-clause which explicates the content of *sum andgit*. This is a type of CLAN ("Clause-and-Nominal", see Warner 1982), where a relatively vague nominal element in one clause has its semantics spelled out in a second clause that is syntactically quite loosely constructed; this is a striking feature of almost all of these *tocnawan*-clauses in Ælfric.

(12)   Subject – Object – **tocnawan** – Vfinite (*magon*) – Postfield
    We wyllað eow secgan sume swutelunge   nu   be   þam halgum
    we will   you tell   some explanation now about the   holy
    tidum   ðe   we healdað and weorðiað on geleaffullum cyrcum mid
    feastdays that we hold   and celebrate in orthodox   church with
    *Godes lofsangum þæt   ge   sum andgit   þæron   **tocnawan***
    God's psalms   so.that you some understanding about.that acquire
    magon, hu   eall ðæs geares ymbegang Gode ælmihtigum ðeowað.
    may,   how all the year's round   God Almighty   serves
    'We will tell you now some explanation about the holy feastdays that we observe and celebrate in the Christian church with God's psalms, so that you may acquire some understanding on that, i.e. how the entire round of the year serves God Almighty.'   <ÆHom 11, 1>

An example more central to the focus of this paper is (13), where we have a main clause that starts with an adverbial containing a link to the previous discourse:

(13)   Adverbial – Subject – Vfinite (*mæg*) – Object – **tocnawan** – Postfield
    *Be ðisum þeawum man mæg þæne man **tocnawan**, hwæðer him*
    by these habits   one may the   man recognize whether him
    godes gast on wunige oððe þæs   gramlican deofles.   <ÆSpir 64>
    God's spirit in dwells or   of.the cruel   devil
    'By this conduct one may distinguish a man who is inhabited by God's spirit from a man who is inhabited by the spirit of a cruel devil.'

The vague NP *þæne man* is further defined by a following clause that is only loosely connected; PDE would be more likely to have (i) a relative clause as in the PDE translation, or (ii) a more tightly constructed indirect question, as in *By this conduct one may tell whether a man is inhabited by God's spirit or by the spirit of a cruel devil*, or of course a "permissive subject" and a flexible-valency verb like *show*, as in *This conduct shows whether a man is inhabited by God's spirit or by the spirit of a cruel devil* (see, e.g., (7a–b) above). Many if not all of the *tocnawan*-examples in this section have the same generic subjects as Alfred's translations in (9)–(11): *By this, **we/ people/you/one** may perceive…* as well as passives like *By this it **may be perceived***, both of which very readily allow rephrasing with a PDE *show* and a "permissive subject" *This* that either links back or links forward, as in *This shows that (…).*

   The set of generic expressions includes the form *man*, as in (13). This *man* (the first occurrence of *man* in the example) is etymologically the same form as the second *man* 'man', but it is a grammaticalized form, no longer a noun but a pronoun (see, e.g., van Bergen 2003) and an ultra-indefinite pronoun at that, only available in subject form, and discursively inert (Los 2002); in sum, a generic placeholder. Statements with generic *man* can be reworded by passives in OE (as is also the case for Dutch *men* and German *man*); and passives are in fact the best option in PDE, given that PDE *one* has been shown not to be an adequate replacement of OE *man* (see Seoane Posse 2000; Los 2002). Ælfric uses the alternation between passives and *man* to vary the wording of his exegetical explanations: *by this we may recognize/by this one may recognize/by this is recognized…* An example of *tocnawan* in the passive is (14):

(14)   Adverbial – Vfinite (*beoð*) – Subject – **tocanawan** – Postfield
       *On ðam beoð cristene   men **tocnawene** gif hi   rihtlice   cristene*
       by that are   Christian men recognized if they truly   Christians
       beoð.   <ÆCHom II, 40 300.30>
       are
       'By that Christians are recognized as true Christians.'

What example (13) above also shows is how finite verb movement in OE is different from V2 in Dutch and German: two constituents precede the finite verb.[2] Such V3 constructions are the norm if the first constituent is something other than a *wh*-constituent, a negated constituent, or one of a small set of adverbs, most prominently *þa* 'then,' and the subject is a pronoun rather than a full NP; this insight is due to van Kemenade (1987). The V3 order of example (13) is in fact a diagnostic

---

2.   Note that I use the term *constituent* as an equivalent of XP (a phrase) as opposed to X (a head); the finite verb is a head (V) and has moved to another head-position ("head-to-head movement"). Any mention of *constituent* in this paper, then, excludes the moved finite verb.

for the pronominal rather than nominal status of *man* in OE (van Bergen 2003). In contrast, Example (14), with its nominal rather than pronominal subject, has only a single constituent before the finite verb. Note that the verb itself is supposed to have moved to the same position in (13) and (14), namely, the lower landing site (see, e.g., van Kemenade 2000; Haeberli 2002). OE has more positions for subjects than Dutch and German, and this is one of the reasons why OE can accommodate the requirements of information structure so readily.

I have argued elsewhere (Los 2012) that the original motivation for the lower landing site may have been information-structural rather than syntactic, with the finite verb functioning as a demarcator for given and new information domains.[3] Note that in (13), the adverbial and the pronoun are "given" information; personal pronouns are "given" by definition. That movement to the lower landing site still has this purely information-structural function in OE is unlikely. We have already seen that the referential adverbials in first position can be forward-looking (like *þi* in (10c)) as well as backward-looking (like *forðæm* in (9c)). The asymmetry between the positions of pronominal and nominal subjects with respect to the verb, as found in (13) and (14), might have been reanalyzed at a later point as syntactic, whereby the higher subject position, the one to the left of the lower landing site, is reserved for pronouns and the one on the right, originally for "new" subjects which naturally would not be pronouns but full nominals, is a position for full nominals. However, there are many examples of nominal subjects to the left of the finite verb (Koopman 1998; Haeberli 2002); there is evidence that such nominal subjects in the higher, pronominal position tend to be both given and specific (e.g. Biberauer & van Kemenade 2013), which might indicate that finite verb movement to the lower position is still discourse-sensitive in (some varieties of) OE. Dreschler's (2015) investigation of nominal subjects after clause-initial PPs gives quite high rates for nominal subjects in this high pronominal position in PP-initial main clauses in *Orosius* (55.4%), against much lower rates in Ælfric's *Lives of Saints* (12.4%) (Dreschler 2015: 259). The type of PP plays a role, with nominal NPs in the high position being particularly frequent after *after*-PPs, as does the type of verb, with unaccusatives disfavoring such nominal NPs. The information status of the nominal NPs in the high position, which could have been expected to be "given" on the basis of the findings of Biberauer & van Kemenade (2013), turn out to be less clearcut; although there is a clear preference for new subjects to be in postverbal position, there are 16 instances in *Orosius* of new subjects in the high position (Dreschler

---

3.   See also Hinterhölzl & Petrova (2010: 319), although they present a unified scenario in which the two landing sites in OE derive from a single site for finite verb movement in earlier Germanic. The competition for the first position in languages generally makes it more likely that OE represents the original situation with two landing sites, with the verb-second rule in Present-day Dutch and German a simplification.

2015: 261); in 13 of these 16 instances, the subject was a proper name, introducing a new referent or reactivating an earlier referent, as in (15):

(15) *Æfter þæm Pompeius se consul for   on   Numentinas, Ispania*
      After that Pompey the consul marched upon Numentines Spain
      þeode.   <Or 5 2.115.22> (Dreschler 2015: 261)
      people
      'After this Pompey, the consul, marched upon the Numentines, a people of Spain.'

The other two examples Dreschler provides to illustrate this pattern start with *On þæm dægum* 'in those days' and *On þære tide* 'at that time', which could suggest that the OE translator of the *Orosius* used the PP – full NP – Vfinite order to mark an episode boundary. This matter has to be left to future research.

The schema in (16) shows that OE syntax makes at least four positions available for subjects. *V* stands for the non-finite verb, if present, and for the original position of the finite verb *v*; "…" stands for the base positions for adverbials, complements/objects to the left of V, which will be spelled out in more detail in (17).

(16) [Prefield: Subject$_1$ – Subject$_2$ – *v*][Middle field: Subject$_3$ – … – V][Postfield: Subject$_4$]

The prefield has two positions, both derived (it could be argued that the prefield is itself created by movement of the finite verb); "Subject$_1$" labels the clause-initial position to which any clausal constituent may move in a V2 language, including the subject.[4] "Subject$_2$" labels the "high" position to the left of the lower landing site of the finite verb *v*. The middle-field position could be regarded as basic, whereas the position in the postfield is the so-called "late subject" position (see Warner 2007).[5] This clausal architecture provides two positions for subjects that encode new information to be positioned towards the end of the clause, after the finite verb, and shows the flexibility of the system.

The options for positioning adverbials is given in (17):

(17) [Prefield: Adverbial$_1$ – … – *v*][Middle field: … – Adverbial$_2$ – … V][Postfield: Adverbial$_3$]

---

4. I am following here the standard V2 analysis in which main-clause-initial subjects are treated on a par with other main-clause-initial constituents, i.e. as derived XPs in Spec,CP, rather than the alternative analysis offered in Travis (1984) in which they are assumed to be in Spec,IP (a critique of this analysis can be found in Koster 2008).

5. I am abstracting away from whether these subjects end up "late" by a Heavy NP Shift-like operation (to a position adjoined to the VP) or whether the subject remains low in the structure, which is a possibility for the subjects of unaccusative verbs, as they start out as internal objects; see again Warner (2007) or van Kemenade (1997).

Although PDE can be argued to have a clause-initial, medial, and clause-final position for adverbials, too, there are important differences in what type of adverbials the PDE positions may accommodate. The default position in PDE is clause-final, where adverbials can be limitlessly stacked; the default position in Dutch and German, and by extension in OE, is the middle-field position, again with limitless stacking. The medial position in PDE is restricted to adverbs rather than phrases, particularly manner adverbs, and stacking is very limited.

The clause-initial position is our main concern in this paper. Of all adverbials, time adverbials are most likely to appear clause-initially in PDE (see Figure 3.3 in Hasselgård 2010: 56), probably because they have an additional text structuring function (see, e.g., Virtanen 1992). Compared to place adverbials, the frequency with which they appear in the initial position is much higher: although, like place adverbials, time adverbials prefer the final position (55% according to Biber et al. 1999: 802), 20% appear in initial and 25% in medial position. Pre-subject adverbials of place are more circumscribed. Biber et al. consider them "very marked" (Biber et al. 1999: 803), and they occur overwhelmingly clause-finally (90%, compared to 5% initially and 5% medially). Their primary function in initial position has been described as framesetting (Chafe 1976, quoted in Krifka 2007: 45), as in (18); the adverbials are given in italics:

(18)   A: How is business going for Daimler-Chrysler?
        B: [*In GERmany*]$_{Frame}$ the prospects are [GOOD]$_{Focus}$, but [*in AMErica*]$_{Frame}$ they are [losing MOney]$_{Focus}$.   (Krifka 2007: 45)

Framesetting sets up the background that limits the scope of a proposition, hence the sense of contrast ("here, but not there"). The important thing to note is that frame-setters do not contain a reference to the preceding discourse. Such non-referential frame-setters are also perfectly possible in OE, in that OE has forward-looking referential clause-initial adverbials, as we have seen already in (10c); an example with *tocnawan* is (19) below:

(19)   Adverbial – Vfinite (***tocnawan***) – Subject – Postfield:
        Be ðam **oncnawað** *ealle men þæt ge   sind mine folgeras   gif ge*
        by that perceive   all   men that you are   my   followers if you
        habbað lufe eow betwynan;   <ÆCHomII, 40 300.32>
        have   love you between
        'All people will recognize that you are my followers if you show love for each other;'

These are not frame-setters but part of the textual coherence strategies so typical of OE in which clauses are syntactically coordinated rather than subordinate, and linked to each other by means of correlative deictic elements.

In contrast, *Be þam* 'by that' in example (20) below is again backward-looking (*þam* refers to Christ's willingness to be born in a human body in order to suffer for the sake of Man's redemption, which is the content of the previous sentence). This example also shows the flexibility of OE in moving constituents like the object-NP *Cristes eadmodnysse* to the postfield, where it receives special prominence. The NP is not particularly heavy in the sense that it contains a lengthy postmodifier like a long PP or a relative clause, which would be required as a condition for moving NPs in Heavy NP Shift in PDE, but moving constituents to the postfield for reasons of prominence rather than heaviness is a feature of Ælfric's style. Note that, once more, we have a CLAN-construction: the NP is followed by a clause which serves to further explain in what way Christ was humble:

(20)   Adverbial – Subject – Vfinite (*magon*) – V (***tocnawan***) – [$_{postfield}$Object – Postfield]:
Be þam we magon ***tocnawan*** *Cristes   eadmodnysse, þæt se healica*
by that we may   perceive   Christ's humility,   that the sublime
God hine sylfne swa geeadmette, þæt he ðam deaðe underhnah and
God him self   so humbled   that he the death suffered   and
þone deofol oferswyðde…   <ÆLS (Memory of Saints) 113>
the   devil overcame
'By that we may perceive Christ's humility, that the sublime God so humbled himself that he allowed himself to suffer death, and overcame the devil.'

Adverbial discourse links are not restricted to the prefield but may freely occur in the other two adverbial positions: in the middle field, like *be þysum* 'by these things' in (21), and in the postfield, like *be þam hæðenum godum* in (22):

(21)   Subject – Vfinite (*magon*) – [middle-field Adverbial] – V (***tocnawan***) – Postfield[6]
We magon be þysum[6] ***tocnawan*** *þæt se mann, þe   his gesyhðe*
we may   by these   perceive   that the man   who his sight
næfð, ne sceal he gedyrstlæcan, þæt he mæssige   þonne he
not.has not shall he presume   that he celebrates.mass when he
ne gesyhð hwæt he offrað Gode, hwæðer þe clæne, þe ful.
not sees   what he offers God, whether or clean   or foul
'We may by these things perceive that the man who has lost his sight should not presume to celebrate mass, as he cannot see whether what he is offering to God is clean or foul.'   <ÆLET 1 (Wulfsige CCCC 190) 148>

Note that (22) has two adverbial discourse links, of which *her* 'here' occupies the clause-initial position while the *be*-PP is in the middle field:

---

6. Or: 'by this' (dative singular). The passage in its entirety allows both interpretations.

(22)   Adverbial – Subject – Vfinite – **tocnawan** – [postfield Adverbial – Postfield]
*Her   we   magon   **tocnawan**   be   þam   hæðenum   godum,   hwilce*
here   we   may   perceive   about   the   pagan   gods   which
mihte   hi   hæfdon   ongean   þone   ælmihtigan   god.   <ÆHom 22, 286>
power   they   had   against   the   Almighty   God
'Here we may perceive how much power the pagan gods had against Almighty God.'

Both (21) and (22) are, yet again, CLAN-constructions: s*e mann … he*; *be þam hæðenum godum … hi*.

The final example shows two main clauses constructed along parallel lines. The first one comprises the verb *understandan*, and the second with *tocnawan*. They are synonyms here, both meaning 'understand, interpret' (*we should interpret Abraham as the Almighty Father… and we should interpret the sacrifice of Isaac as the Lord's passion*); the relevant adverbials appear in italics.

(23)   Subject – Vfinite (*sceolon*) – **understandan/tocnowan** – [postfield Adverbial – Object – …]
*We   sceolon   understandan   **on abrahame**   þone   ælmihtigan   fæder   and*
we   should   understand   by Abraham   the   Almighty   Father   and
***on Isaace**   his   leofan   sunu   urne   hælend   crist.*
by Isaac   his   dear   son   our   Saviour   Christ
***be   ðam**   cwæð   se   heofonlica   fæder.*
about that.one   spoke   the   heavenly   father
þes   is   min   leofa   sunu   ðe   me   wel   licað.
this   is   my   dear   son   who   me   well   pleases
*and   we   sceolon   tocnawan   **on isaaces offrunge**   drihtnes   ðrowunge.*
and   we   should   understand   by Isaac's   sacrifice   Lord's   passion
***Be   ðam**   cwæð   se   apostol   paulus.*
about that.one   spoke   the   apostle   Paul
þæt   god   fæder   ne   sparode   his   agenum   bearne.
that   good   father   not   spared   his   own   child
ac   for   us   eallum   hine   to   deaðe   sealde   <ÆCHom II, 4 34.161>
but   for   us   all   him   to   death   gave
'We should interpret Abraham as the Almighty Father and Isaac his dear son, as our Savior Christ. About him the heavenly father said: this is my dear son, who pleases me well. And we should interpret the sacrifice of Isaac as the Lord's passion. About that [sacrifice] the apostle Paul said: the good father did not spare his own child but gave him up to death for our sakes.'

Each of these clauses is followed by a clause with an adverbial link (*Be ðam* 'by that'). Such parallel structures – including the echoing *Be ðam cwæð se heofonlica fæder/Be ðam cwæð se apostol paulus* – are very typical of Ælfric's polished style.

These clause-initial *be* 'about'-PPs cannot be easily translated into idiomatic PDE. In the first PP, the demonstrative needs to be translated in PDE by a personal pronoun ('about him'), as singular independent demonstratives can no longer be used to refer to people (see, e.g., Huddleston & Pullum 2002: 1504); in the second PP, the demonstrative does not establish the referential relationship with the sacrifice of Isaac in the preceding clause as straightforwardly as in PDE – idiomatic translations would spell such relationships out more explicitly: 'about this sacrifice'. Furthermore, clause-initial adverbials have become marked in PDE and are no longer unproblematically available as links to the immediate discourse.

Unlike preposed objects, clause-initial adverbials show a decline in referentiality over time in that they tend to encode new information more often in PDE than in EModE or LME; quantitative evidence for this can be found in Pérez-Guerra (2005: 358), Los (2012), Komen et al. (2014), and Dreschler (2015: 242–265, 300–317). If clause-initial adverbials are less often referring (either forwards or backwards in PDE (see (18)), a loss in referentiality in these PPs over the years is just what we would expect.

To summarize: what the OE examples of *tocnawan* 'understand' show is that verbs of thinking and declaring have predominantly personal subjects in OE, even in argumentative prose where such subjects are likely to be generic rather than specific ("we", "you", "people in general"). PDE would go for a much more impersonal style with non-agentive subjects, along the lines of constructions like *This parabel shows/This text should be interpreted as*. Such non-agentive subjects are likely to be discourse links, following a line of argument. The OE examples use adverbials to establish such links, and these adverbials are facilitated by the flexible clausal architecture of OE, where pre- and middle-field adverbial positions are not earmarked for specific adverbials, unlike clause-initial and clause-medial positions in PDE.

With the decline of adverbial discourse linking, and subjects taking over as the expression of choice for discourse links, there was a need for strategies to allow discourse links – whether animate or inanimate – to be expressed as subjects. One such strategy was new passives, like the passive ECM-construction with verbs of thinking and declaring that we discussed in the previous section, or passives like *He was prescribed lithium* which added RECIPIENTS to the thematic roles that can appear as subject (see Dreschler 2015 for a historical survey of "new" passives). In addition to passives, which are clearly signaled by a formal construction, there were other strategies that involved a change in argument structure without formal marking. They are the topic of the next section.

4.   The rise of valency alternations

This section discusses two further strategies that promote "permissive subjects": middles, like *This book reads well*, and verbs with double argument structures, either causative/ergative pairs as *in John broke the vase/The vase broke* or the locative alternation seen in pairs like *Jill sprayed paint on the wall/Jill sprayed the wall with paint* (see, e.g., Levin 2006). A third development that led to new strategies, namely, impersonal verbs (*Me likes*, cf. *me … licað* in (23)) becoming personal verbs (*I like*), resulted in EXPERIENCERS being expressed by "proper" nominative subjects rather than by dative or accusative NPs. This development has been discussed extensively in the literature (see Allen 1995; Möhlig-Falke 2012) and will not be discussed here.

The rise of middles in the history of English is discussed in Hundt (2007). Middles are often associated with passives in that they perform a similar function, i.e. promoting an object (with the thematic role of PATIENT/THEME) to subject. Middles involve the intransitive use of a transitive verb. There are some genres in PDE that are particularly permissive in creating middles from scratch, as it were, which shows that the pattern has acquired a dynamics of its own, allowing entire sets of verbs to be used as middles. One such genre is mail order catalogues, where Hundt found examples such as the following:

(24)   Matching hood *converts* into collar.

(Sears & Roebuck Catalogue) (Hundt 2007: 161)

Middles are a special case of causative/ergative valency alternations. Valency alternations are crosslinguistically common for a small set of verbs, typically with change-of-state meanings like *break, burn, open/close, begin/stop/continue*. Härtl (2003) identifies a number of conditions that need to be met for such verbs to occur as intransitives in German. For one, the THEME/PATIENT argument must possess inherent properties that allow the change of state to proceed independently, without an agent. This explains why *zerreißen* 'tear apart' can be used intransitively with entities such as a *sail*, but not with a *picture* (Härtl 2003: 909). Conversely, transitive uses are impossible if the change of state cannot be conceptualized (in terms of "naïve physical reasoning" (907)) as being initiated by an AGENT: *verrosten* 'rust' or *verwittern* 'weather' can only be used intransitively. These conditions account for the alternating argument structures in English verbs like *break, burn, open/close, begin/stop/continue*, where both conceptualizations are possible. What is special about PDE is that entire classes of verbs partake in these alternations. Any de-adjectival verb, like *blacken, pretty up*, or *liquify* participates (Francis & Sinclair 1994), as do many "manner of motion" verbs:

(25) a. Water *seeped* through the roof of the tunnel.
      b. The roof of the tunnel was *seeping* water.
                        (Hawkins 1986: 58–61, from Rohdenburg 1974)

(26) a. The car *nosed* into the city traffic
      b. I *nosed* the car onto the tracks.   (Francis & Sinclair 1994: 198)

Consider also the more extreme cases below, with *post, ship*, and *see*:

(27) … this is the second time I've written this out. The first time – blogger messed
      up on me and it didn't *post* and I lost the whole thing.
                        (Blogcorpus, 2059313.male.26.Student.leo.xml, Schler et al. 2006)

(28) Your order has *shipped*.   (Amazon.com)

These intransitive uses violate the condition identified by Härtl that the entity must
have an inherent property that allows the process to go ahead without an AGENT.
See in (29) is a transitive verb, but its subject is not the EXPERIENCER that would be
expected on the basis of the conceptualization of a seeing-event:

(29) 2012 *saw* the second highest carbon emissions in half a century.
              (Colin Schultz, smithsonian.com, March 7, 2013; accessed 18 October 18 2015)

A relevant example from the contrastive English/Portuguese pairs in Santos (1988)
is (30):

(30) a. he pays – *ele **paga***
      b. crime pays – *o crime **compensa***

The rise of "permissive" valencies as in (24)–(26) has been investigated by Van
Gelderen (2011). She shows that these flexible valency alternations are a relatively re-
cent phenomenon; they are the result of two processes: transitive verbs acquiring in-
transitive counterparts, like the middle constructions in (24), and intransitive verbs
acquiring transitive counterparts, as in the manner of motion verbs in (25)–(26).
     In (24), (26a), (27), and (28), the "permissive subjects" appear to encode PATIENTS
or THEMES (a hood, a car, a blogpost, and a delivery order), but in (25b) we have a
SPACE, in (29) a TIME, and in (30b) a SOURCE. These uses go well beyond providing
an alternative expression for discourse-linking adverbials; only (25b) and (29) can be
argued to have a clause-initial PP-alternative (*Through the roof of the tunnel*; *In 2012*).
     Van Gelderen (2011) notes that what seems to have caused the initial push may
have been various losses in derivational morphology. There is a core set of verbs
whose double argument structure is of long standing, like the causative/unaccusa-
tive pair in (31a–b):

(31) a. The ship *sank*.
      b. The admiral *sank* the ship.

These verbs can be traced back to Proto-Germanic causatives built on o-grade stems of intransitive verbs – both unaccusatives and unergatives – by means of a *-ja-* suffix. An example is **siŋkwana* 'sink' (intr.), corresponding to *sink* in (31a), which gave rise to causative **saŋkwjana* 'sink' (tr.), corresponding to *sink* in (31b). Some of these pairs still have distinct forms in PDE (data from Ringe 2006: 253–254):

(32) a. **ligjana* 'lie' (intr.) and **lagjana* 'lay' (tr.)
      b. **rīsana* 'rise' (intr.) and **raizijana* 'raise/rear' (tr.)
      c. **sitjana* 'sit' (intr.) and **satjana* 'seat, set' (tr.)

Many other pairs lost one of their members, with the remaining one taking over for both. For instance, PDE *sink* in (31a) would have been OE *sincan*, while PDE *sink* in (31b) had its own derived OE form *sencan*, still surviving in ME as *sink/sench*. These mergers of intransitives with their causative counterparts are not specific to English but a general Germanic phenomenon; there are non-standard varieties of PDE and Dutch where *lie/lay* have merged to *lay* (Dutch *leggen*). One vector for such mergers is the fact that the perfects of unaccusatives, which select *be*, as in (33), cannot be distinguished from the passive of their causative counterpart:

(33) *þa   wæs an gereord   on   eallum mancynne, and þæt weorc **wæs***
      then was one language among all   mankind   and the work was
      ***begunnen** ongean godes willan*;   <ÆCHom I, 22 318.17>
      begun   against god's will
      'At that time all people spoke one language, and the work had been begun/was begun against God's will.'

Van Gelderen's (2011) investigation finds that OE already had quite a considerable number of transitive/intransitive ("labile") verbs compared to Dutch and German, and that this number has been increasing ever since: 223 OE exclusively intransitive verbs are reduced to only 30–40 in Modern English, while 80 "labile" verbs in OE increase to about 800 in PDE (van Gelderen 2011: 119, 122).

What is specific to English is the phenomenon of further mergers as a result of the loss of verbal prefixes. Van Gelderen particularly mentions the prefix *ge-*, and there are others. Consider the following much quoted example of a valency change without derivational morphology:

(34) a. They *loaded* hay onto the wagon.
      b. They *loaded* the wagon with hay.

The locative alternation of (34a–b) requires a formal contrast in Dutch and German, with (34a) using a simplex verb and (34b) a prefixed verb (usually with the prefix

be-) (Brinkmann 1997); note that in this particular case, the prefix *vol-* is also a distinct possibility as a Dutch correlate of (34b) (p.c. Hendrik de Smet). The increases in labile verbs have apparently over time acquired a dynamics of their own, allowing entire classes of verbs to appear with such subjects. The proliferation of "permissive subjects" may well outstrip the decline in referential adverbials linking to the previous discourse, so that the rise of subjects as the discourse link of choice cannot be expected to be a straightforward case of competition.

5.   Adverbial discourse links, "late subjects", and flexible argument structure in PDE

Adverbial discourse links do survive in one particular niche in PDE: subject-verb inversion, which should probably be analyzed as the "late subject" construction rather than movement of the finite verb (for "late subjects" in the history of English, see Warner 2007). The information structure of these PDE instances has been investigated in Birner & Ward (1998). The construction has a discourse link as first constituent (either an adverbial as in (35) or a subject complement, in this case a passive participle, as in (36) – both in italics), which allows the subject to be in end-focus position. This is one of the marked constructions in PDE that function as escape hatches for subjects that are not discourse-old, which would compromise the natural flow of information in the canonical SVO order:

(35)   In the iron trade, enormous quantities of material are used for the manufacture of boilers and pipes; while the manufacturers of paint, putty, and other materials also do a brisk trade with market growers. *To these* must be added the various gas companies and colliery merchants, who provide thousands of tons of coke or anthracite coal to feed the furnaces attached to the glasshouses.

[weathers-1913,1,7.154-156][7]

(36)   This blue jug was one of the many objects found in the grave alongside 'Ivory Bangle Lady'. *Also discovered* was an openwork mount of bone with the inscription 'Hail, sister, may you live in God,' indicating Christian beliefs.

(<http://news.bbc.co.uk/local/york/hi/people_and_places/history/
newsid_8537000/8537231.stm>)

There is a connection with flexible argument structure in that corpus studies of subject-verb inversion have shown that they are particularly frequent with intransitives of transitive/intransitive pairs in PDE (as noted by Biber et al. 1999: 954); examples are in (37)–(40) (discourse links in italics, verbs in bold):

---

7.   References in square brackets [ ] are from the *Parsed Corpus of Modern British English* (Kroch et al. 2010).

(37)   The care of his baggage made Paradis divide his men into two bodies, *between which* **marched** the Indians, called Coolies, who carried his chests.

(*OED* 1763 R. Orme Hist. Mil. Trans. Brit. Nation I. 80)

(38)   *Through his translucent skin* **showed** the blue veins, his insides as visible and vulnerable as a tiny transparent shrimp.   (Jane Rogers, *Her Living Image*, Chapter 10; Faber & Faber 1984; e-version Canongate Books 2012)

(39)   *Round her* **burned** iron-spiked circles of tapering candles, yellow-bright in the dark. Before her lay heaps of flowers.   (*BNC*; Biber et al. 1999: 954)

(40)   *Behind her* **trailed** a gaggle of over-age girls-about-town all with their eyes open for the main chance.   (*BNC*; Biber et al. 1999: 954)

This suggest that flexible valency in PDE not only facilitates "permissive subjects", as a way of encoding discourse links, but also plays a role in the positioning of subjects containing new information. In the earlier system, this may not have been a specific function of clause-initial (place) adverbial links. The different functionality of PDE (place) adverbial links is another reason why a decline in referential adverbials linking to the previous discourse and the rise of "permissive subjects" cannot be expected to be a straightforward case of competing structures.

6.   Tail–head linking in Dutch and German narratives

There is clear quantitative evidence in Pérez-Guerra (2005), Komen et al. (2014), and Dreschler (2015) for the decline of adverbial discourse-linking of the type found in OE, as outlined in the previous section, and there is a rise in "permissive subjects". We saw that literal PDE translations of OE texts improve if we express linking adverbials as subjects, which suggests that at least some of the adverbial losses are compensated for by the subject taking over the discourse-linking function. Unlike other syntactic developments, like *have*-perfects taking over from *be*-perfects (Rydén & Brorström 1987; Kytö 1997) or *to*-infinitives encroaching on the domain of subjunctive clauses (Los 2005), this development cannot be supported by straightforward quantitative evidence of S-curves typical of competing structures; see Van de Velde (2014) for other examples of this phenomenon. As we noted in the discussion of examples (24)–(30), the mechanisms that led to entire classes of verbs acquiring flexible argument structures have produced "permissive subjects" that do not have adverbial counterparts in Dutch and German, and by extension, probably not in OE either. There is also evidence that the need for discourse linking itself experienced a decline. The evidence for this comes from contrastive psycholinguistic studies which show that Dutch and German narratives show a system of tail–head linking that has no equivalent in PDE. This tail–head linking

is facilitated by the same type of prefield and middle-field adverbial positions that we discussed in Section 3.

Psycholinguistic studies demonstrate that German speakers show a higher degree of granularity in their descriptions than PDE (Carroll & Lambert 2003). Example (35) shows that German speakers, asked to describe a picture, barge straight in with fine-grained descriptions of where various landmarks are:

(41) a. *Auf den linken Seite ist eine Apothek.*
          on the left   side is a   drugstore
      b. *Vorne   im   Bild   ist eine Strasse.*
          at.the.front in.the picture   is a   street
                                      (Carroll & Lambert 2003: 269)

English informants, on the other hand, tend to establish the global topic first – the picture – and then go on to point out the existence of landmarks by using existential *there*, rather than specific place adverbials:

(42) This is a picture of a busy square.
      There is a square with a fountain.   (Carroll & Lambert 2003: 269)

This difference between the two groups of speakers suggests that the PDE speakers are less likely to use place adverbials. This is further supported by descriptions of events. An experiment described in Carroll et al. (2004) involved English and German retellings of episodes from a short narrative (a clay animation film *The Quest*). Following the procedure that gave rise to the corpus of PDE retellings, a Dutch corpus was constructed by van Ierland (2009); this Dutch corpus was further extended at the Radboud University Nijmegen by Bouwmans (2009) and myself.[8] The PDE and the Dutch retellings reveal a marked difference in the use of adverbial linking which shows that the Dutch informants use place adverbials much more frequently, even when retelling the same events. The different rates of adverbial usage between the two groups hinges on second mentions.

Both groups mention places as adverbials when such places become relevant in the narrative events. The examples in (43a)–(44a) show that both groups mention a grid when it first heaves into view, as an object on which the clay man, the protagonist of the story, finds himself lying after falling on his face. Once the grid has been mentioned (as in (43a)) and is presumably totally activated and in focus in the

---

8.   The Dutch corpus consists of data collected by Suzan van Ierland (van Ierland 2009) with additional data collected by the author and Justine Bouwmans (Bouwmans 2009). All data were collected at Radboud University Nijmegen (53 retellings in all). The English data corpus used here is the one collected by Suzan van Ierland at Birkbeck College, London, with English (British and American) informants (van Ierland 2009: 206), augmented by data collected in Nijmegen (21 retellings in all).

hearers' minds, the English informants tend to produce utterances like (43b1–3) which do not make explicit mention of the grid, even though it plays a part in what happens next: the clay man looks through it and spots running water. The single exception in the corpus is (43b4):

(43) a. he falls onto a kind of grid…
    b. 1. he looks down
       2. he peers through below
       3. he can see way down to the bottom
       4. he can see the water down below *through the grid*

In contrast, the majority of the Dutch informants refer back to it in their b-utterances, using the prefield:

(44) a. *Hij valt met zijn buik op een soort rooster.*
      he falls with his belly onto a kind.of grid
    b. **Door dat rooster** ziet hij beneden water stromen.
      through that grid sees he below water flow

There is a similar example of an instrument, a piece of rock that the protagonist uses to knock a hole in a boulder. Both groups mention the protagonist finding a piece of rock and picking it up (45a–46a), but the English informants leave it implicit in the b-utterance that describes its use:

(45) a. so he finds another rock that's kind of smaller than this than this one, and he picks it up
    b. and he starts beating this rock, he starts trying to break this rock open

The Dutch informants mention the rock explicitly when they describe the beating action, using a pronominal adverb in the prefield adverbial position in (46a) or the middle-field position in (46b):

(46) a. *pakt een stuk rots,* **daarmee** *slaat hij op de waterplek*
      takes a piece rock, there.with hits he on the water.spot
      '(he) takes a piece of rock, with this he hits the water spot'
    b. *hij pakt een rots en slaat* **ermee** *op de steen waar het*
      he takes a rock and hits it.with on the boulder where the
      water valt
      water falls
      'he takes a rock and hits with it on the boulder where the water is dripping'

The difference between the two groups can be described in terms of tail–head linking, a narrative technique in which each new move in a narrative is explicitly linked to what has gone before by means of repetition. If Old English is like Dutch

and German, one of the responses to the decline in adverbial linking can have been less tail–head linking, and a more implicit descriptive style, which does not require place adverbials.

If this type of tail–head linking itself declines, this is another reason why a decline in referential adverbials linking to the previous discourse and the rise of "permissive subjects" cannot be expected to be a straightforward case of competing structures.

7.   Conclusions

This paper has presented a survey of a number of phenomena that have been reported in the literature on the history of English and in contrastive crosslinguistic studies and links them in a scenario of change in which the functionality of clause-initial adverbials as linkers to the immediately preceding discourse has been lost, and whereby the role of discourse linking of the subject has been further extended. As the subject has become the unmarked way to express discourse links in PDE, the functional load of the subject has increased. Discourse links are often non-agentive inanimates, and these are readily accommodated in PDE by a range of strategies for creating subjects: more options for passivization, and productive, flexible valency patterns; both are developments that have given rise to crosslinguistically-unusual "permissive subjects".

The best evidence to support a link between a decline in one structure and a rise in another can be provided by a quantitative investigation, but the difficulty is that there is unlikely to be quantative evidence of a direct competition between discourse links expressed by adverbials and by subjects.

i.   The flexibility of PDE verbs to acquire additional argument structure (intransitives developing a transitive use, transitives developing an intransitive use) has acquired a dynamics of its own, which is likely to have led to a proliferation of "permissive subjects" far beyond what might have been needed to compensate for the loss in referential adverbials (Sections 2 and 4).

ii.   The referential adverbials that are left seem to occur in the "late subject" construction (what has traditionally been called "locative inversion" or "subject-verb inversion" in PDE), which has the primary function of presenting new information (see Section 5); this does not appear to be the predominant function of clause-initial adverbials in OE, as we saw in Section 3.

iii.   Adverbial links are no longer required as a mechanism required for tail–head linking in narratives, as PDE narratives no longer appear to involve tail–head linking at all (see Section 6); this would mean that the need for discourse linking in narratives has itself decreased.

What is a matter of future research is a more detailed investigation into OE style, and a more fine-grained investigation of the use of the subject in narratives in the history of English as the expression of protagonists, non-protagonists and discourse links.

References

Allen, Cynthia L. 1995. *Case-marking and Reanalysis: Grammatical Relations from Old to Early Modern English*. Oxford: Clarendon Press.

Bergen, Linda van. 2003. *Pronouns and Word Order in Old English, with Particular Reference to the Indefinite Pronoun man*. London: Routledge.

Biber, Douglas, Johannsson, Stig, Leech, Geoffrey, Conrad, Susan & Finnegan, Edward. 1999. *The Longman Grammar of Spoken and Written English*. Harlow: Longman.

Biberauer, Theresa & van Kemenade, Ans. 2013. Subject positions and information-structural diversification in the history of English. *Catalan Journal of Linguistics* 10: 17–69.

Birner, Betty & Ward, Gregory. 1998. *Information Status and Canonical Word Order in English* [Studies in Language Companion Series 40]. Amsterdam: John Benjamins. doi:10.1075/slcs.40

Birner, Betty & Ward, Gregory. 2002. Information packaging. In *The Cambridge Grammar of the English Language*, Rodney Huddleston & Geoffrey K. Pullum (eds), 1363–447. Cambridge: CUP.

Bouwmans, Justine. 2009. Het zandmannetje is walking naar het water: The spoken description of narratives by Dutch and English speakers and by Dutch students of English. MA thesis, Radboud University Nijmegen.

Brinkmann, Ursula. 1997. *The Locative Alternation in German: Its Structure and Acquisition* [Language Acquisition and Language Disorders 15]. Amsterdam: John Benjamins. doi:10.1075/lald.15

BNC (The British National Corpus), version 3, BNC XML Edition. 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. <http://www.natcorp.ox.ac.uk/>

Carroll, Mary & Lambert, Monique. 2003. Information structure in narratives and the role of grammaticised knowledge: A study of adult French and German learners of English. In *Information Structure and the Dynamics of Language Acquisition* [Studies in Bilingualism 26], Christine Dimroth & Marianne Starren (eds), 267–87. Amsterdam: John Benjamins. doi:10.1075/sibil.26.13car

Carroll, Mary, von Stutterheim, Christiane & Nuese, Ralph. 2004. The language and thought debate: A psycholinguistic approach. In *Multidisciplinary Approaches to Language Production*, Thomas Pechmann & Christopher Habel (eds), 183–218. Berlin: Mouton de Gruyter. doi:10.1515/9783110894028.183

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In *Subject and Topic*, Charles N. Li (ed.), 27–55. New York NY: Academic Press.

Downing, Angela & Locke, Philip. 1995. *A University Course in English Grammar*. Hemel Hampstead: Phoenix ELT.

Dreschler, Gea. 2015. *Passives and the Loss of Verb Second: A Study of Syntactic and Information-structural Factors* [LOT Dissertations in Linguistics 402]. Utrecht: LOT.

Fanego, Teresa. 1992. *Infinitive Complements in Shakespeare's English*. Santiago de Compostela: Universidade de Santiago de Compostela (Servicio de Publicacións e Intercambio Científico).

Fischer, Olga. 1989. The origin and spread of the accusative and infinitive construction in English. *Folia Linguistica Historica* 8: 143–217.

Francis, Gill & Sinclair, John McHardy. 1994. I bet he drinks Carling Black Label: A riposte to Owen on corpus grammar. *Applied Linguistics* 15: 190–200. doi:10.1093/applin/15.2.190

Gelderen, Elly van. 2011. Valency changes in the history of English. *Journal of Historical Linguistics* 1(1): 106–143. doi:10.1075/jhl.1.1.05van

Haeberli, Eric. 2002. Observations on the loss of verb second in the history of English. In *Studies in Comparative Germanic syntax: Proceedings from the 15th Workshop on Comparative Germanic Syntax* [Linguistik Aktuell/Linguistics Today 53], C. Jan-Wouter Zwart & Werner Abraham (eds), 245–72. Amsterdam: John Benjamins. doi:10.1075/la.53.15hae

Halliday, Michael A. K. 1967. Notes on transitivity and theme in English (parts I and II). *Journal of Linguistics* 3: 37–81. doi:10.1017/S0022226700012949

Halliday, Michael A. K. 1968. Notes on transitivity and theme in English (Part III), *Journal of Linguistics* 4: 153–308. doi:10.1017/S0022226700001882

Härtl, Holden. 2003. Conceptual and grammatical characteristics of argument alternations: The case of decausative verbs. *Linguistics* 41(5): 883–916. doi:10.1515/ling.2003.029

Hasselgård, Hilde. 2010. *Adjunct Adverbials in English*. Cambridge: CUP. doi:10.1017/CBO9780511676253

Hawkins, John. 1986. *A Comparative Typology of English and German*. London: Croom Helm.

Healey, Antonette diPaolo & Venezky, Richard L. 1985[1980]. *A Microfiche Concordance to Old English*. Toronto: Pontifical Institute of Mediaeval Studies.

Hinterhölzl, Roland & Petrova, Svetlana. 2010. From V1 to V2 in West Germanic. *Lingua* 120: 315–328. doi:10.1016/j.lingua.2008.10.007

Huddleston, Rodney & Pullum, Geoffrey K.. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.

Hundt, Marianne. 2007. *English Mediopassive Constructions: A Cognitive, Corpus-based Study of their Origin, Spread, and Current Status*. Amsterdam: Rodopi. doi:10.1163/9789401203784

Ierland, Suzan van. 2009. *Grammatical Features Influencing Information Structure: The Case of L1 and L2 Dutch and English*. Utrecht: Lot Dissertations Series.

James, H. R. 1897. *Consolatio Philosophiae*. London. Project Gutenburg, HTML conversion.

Kemenade, Ans van. 1987. *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.

Kemenade, Ans van. 1997. V2 and embedded topicalisation in Old and Middle English. In *Parameters of Morphosyntactic Change*, Ans van Kemenade & Nigel Vincent (eds), 326–352. Cambridge: CUP.

Kemenade, Ans van. 2000. Jespersen's cycle revisited: Formal properties of grammaticalization. In *Diachronic Syntax: Models and Mechanisms*, Susan Pintzuk, George Tsoulas & Anthony Warner (eds), 51–74. Oxford: OUP.

Kemenade, Ans van & Westergaard, Marit. 2012. Syntax and information structure: Verb-second variation in Middle English. In *Information Structure and Syntactic Change in the History of English*, Anneli Meurman-Solin, Maria José López-Couso & Bettelou Los (eds), 87–118. New York NY: OUP. doi:10.1093/acprof:oso/9780199860210.003.0005

Komen, Erwin, Hebing, Rosanne, Kemenade, Ans van & Los, Bettelou. 2014. Quantifying information structure change in English. In *Information Structure and Syntactic Change in Germanic and Romance languages* [Linguistik Aktuell/Linguistics Today 213], Kristin Bech & Kristine Gunn Eide (eds), 81–110. Amsterdam: John Benjamins. doi:10.1075/la.213.04kom

Koopman, Willem. 1998. Inversion after single and multiple topics in Old English. In *Advances in English Historical Linguistics*, Jacek Fisiak & Marcin Krygier (eds), 135–149. Berlin: Mouton de Gruyter.

Koster, Jan. 2008. A reappraisal of classical V2 and scrambling in Dutch. *Groninger Arbeiten zur germanistischen Linguistik* 46: 27–54.

Krifka, Manfred. 2007. Basic notions of information structure. In *The Notions of Information Structure* [Interdisciplinary Studies on Information Structure 6; Working Papers of The Sonderforschungsbereich 632], Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds), 13–55. Potsdam: Potsdam Universitätsverlag.

Kroch, Anthony, Santorini, Beatrice & Diertani, Ariel. 2010. Penn Parsed Corpus of Modern British English. <http://www.ling.upenn.edu/hist-corpora/PPCMBE-RELEASE-1/index.html>

Kytö, Merja. 1997. Be/Have + past participle: The choice of the auxiliary with intransitives from Late Middle to Modern English. In *English in Transition: Corpus-based Studies in Linguistic Variation and Genre Styles*, Matti Rissanen & Merja Kytö (eds), 17–86. Berlin: Mouton de Gruyter. doi:10.1515/9783110811148.17

Levin, Beth. 2006. English object alternations: A unified account. Ms, Stanford University. <http://web.stanford.edu/~bclevin/pubs.html>

Los, Bettelou. 2002. The loss of the indefinite pronoun *man*: Syntactic change and information structure. In *English Historical Syntax and Morphology* [Current Issues in Linguistic Theory 223], Teresa Fanego, María José López-Couso & Javier Pérez-Guerra (eds), 181–202. Amsterdam: John Benjamins. doi:10.1075/cilt.223.12los

Los, Bettelou. 2005. *The Rise of the* To-*Infinitive*. Oxford: OUP. doi:10.1093/acprof:oso/9780199274765.001.0001

Los, Bettelou. 2009. The consequences of the loss of verb-second in English: Information structure and syntax in interaction. *English Language and Linguistics* 13(1): 97–125. doi:10.1017/S1360674308002876

Los, Bettelou. 2012. The loss of verb-second and the switch from bounded to unbounded systems. In *Information Structure and Syntactic Change in the History of English* [Oxford Studies in the History of English], Anneli Meurman-Solin, Maria José López-Couso & Bettelou Los (eds), 21–46. New York NY: OUP. doi:10.1093/acprof:oso/9780199860210.003.0002

Los, Bettelou & Dreschler, Gea. 2012. The loss of local anchoring: From adverbial local anchors to "permissive subjects". In *The Oxford Handbook of the History of English*, Terttu Nevalainen & Elizabeth C. Traugott (eds), 859–871. New York NY: OUP.

Mitchell, Bruce, Ball, Catherine & Cameron, Angus (eds). 1975. Short titles of Old English texts. *Anglo-Saxon England* 4: 207–221. doi:10.1017/S0263675100002799

Mitchell, Bruce, Ball, Catherine & Cameron, Angus (eds). 1979. Addenda and corrigenda. *Anglo-Saxon England* 8: 331–333. doi:10.1017/S0263675100003124

Möhlig-Falke, Ruth. 2012. *The Early English Impersonal Construction: An Analysis of Verbal and Constructional Meaning*. New York NY: OUP. doi:10.1093/acprof:oso/9780199777723.001.0001

OED Online. 2004. Oxford: OUP. <http://www.oed.com/> (30 October 2015).

O'Donnell, James J. (ed.). 1990. *Boethius' Consolatio Philosophiae*, 2nd edn (Bryn Mawr Latin Commentaries). Bryn Mawr PA: Bryn Mawr College.

Pérez-Guerra, Javier. 2005. Word order after the loss of the verb-second constraint or the importance of early Modern English in the fixation of syntactic and informative (un-)markedness. *English Studies* 86: 342–69. doi:10.1080/00138380500164083

Pintzuk, Susan & Taylor, Ann. 2006. The loss of OV-order in the history of English. In *The Handbook of the History of English*, Ans van Kemenade & Bettelou Los (eds), 249–78. Oxford: Blackwell. doi: 10.1002/9780470757048.ch11

Postal, Paul M. 1974. *On Raising: One Rule of English Grammar and its Theoretical Implications*. Cambridge MA: The MIT Press.

Ringe, Don. 2006. From Proto-Indo-European to Proto-Germanic: A linguistic history of English I. Oxford: OUP. doi: 10.1093/acprof:oso/9780199284139.001.0001

Rohdenburg, Gunther. 1974. Sekundäre Subjektivierungen im Englischen und Deutschen: Vergleichende Untersuchungen zur Verb- und Adjektivsyntax [PAKS-Arbeitsbericht Nr. 8]. Bielefeld: Cornelson-Velhagen and Klasing.

Rydén, Mats & Brorström, Sverker. 1987. *The* be/have *Variation with Intransitives in English, with Special Reference to the Late Modern Period* [Acta Universitatis Stockholmiensis 70]. Stockholm: Almqvist & Wiksell.

Santos, Diana. 1988. An MT prototype from English to Portuguese. IBM report. <www.linguateca.pt/Diana/download/Santos88IBMThornwood.pdf>

Schler, Jonathan, Koppel, Moshe, Argamon, Schlomo & Pennebaker, James W. 2006. Effects of age and gender on blogging. In *Computational Approaches to Analyzing Weblogs: Papers from the 2006 AAAI Spring Symposium*, Nicolas Nicolov, Franco Salvetti, Mark Liberman & James H. Martin (eds), 199–205. Menlo Park CA: AAAI Press.

Seoane Posse, Elena. 2000. Impersonalising strategies in early Modern English. *English Studies* 81: 102–116. doi: 10.1076/0013-838X(200003)81:2;1-T;FT102

Taylor, Ann & Pintzuk, Susan. 2012. The effect of information structure on object position in Old English: A pilot study. In *Information Structure and Syntactic Change in the History of English*, Anneli Meurman-Solin, Maria José López-Couso & Bettelou Los (eds), 47–65. New York NY: OUP. doi: 10.1093/acprof:oso/9780199860210.003.0003

Travis, Lisa. 1984. Parameters and Effects of Word Order Variation. PhD dissertation, MIT.

Van de Velde, Freek, 2014. Degeneracy: The maintenance of constructional networks. In *The Extending Scope of Construction Grammar*, Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds), 141–179. Berlin: De Gruyter Mouton.

Virtanen, Tuija. 1992. *Discourse Functions of Adverbial Placement in English: Clause-initial Adverbials of Time and Place in Narratives and Procedural Place Descriptions*. Åbo: Åbo Akademi University Press.

Warner, Anthony R. 1982. *Complementation in Middle English and the Methodology of Historical Syntax*. London: Croom Helm.

Warner, Anthony R. 2007. Parameters of variation between verb-subject and subject-verb order in late Middle English. *English Language and Linguistics* 11: 81–111. doi: 10.1017/S1360674306002127

Cognate noun constructions in Early Modern English

The case of Tyndale's New Testament

Nikolaos Lavidas

National and Kapodistrian University of Athens

This paper examines cognate noun constructions (CNCs) (e.g. *smile a disarming smile*) in Early Modern English, particularly in the first complete English translation of the Bible from the original Greek and Hebrew by William Tyndale. Tyndale's translation is produced during a period of significant expansion of CNCs in English. It is argued that CNCs in Tyndale are a marker of a particular English biblical register which involves archaic (early) English properties (cognate nouns in PPs) rather than a new tendency for cognate direct objects or the result of a translation effect alone. In other words, it is shown that Tyndale's translation follows archaic/early English rules, thus deviating both from the new tendency for cognate direct objects and from the source text. This archaic characteristic of CNCs with cognate nouns in PPs instantiates a general tendency in Tyndale's translation to use archaisms – as evidenced, for instance, in his dispreference of auxiliaries.

Keywords: Early Modern English, transitivity, cognate nouns, biblical English, translation effects, "sacral stamp"

1. Introduction

Cognate noun constructions (CNCs) are constructions in which a transitive verb, or even an intransitive verb, combines with a direct object ("accusative NP")[1] which is a nominalized form of the verb without overt affix marking, or which has the same morphological stem as the verb (Jones 1988: 89). An example is (1):

(1) He *laughed* a loud *laugh*.

---

1. The cognate complement can also appear in the dative case (when the dative is available; Old English) or in a PP; see the discussion of the data below.

In an earlier study (Lavidas 2013a), I showed – in line with Visser (1963–73 [2002]: 415) – that the range of CNCs extended from Late Middle English to Early Modern English to include aspectual CNCs; that is, an increasing number of intransitive verbs with activity/event nouns (see below) were allowed in the construction. In this study, the development of these new CNCs in Early Modern English is examined, and particularly their status in the first complete English translation of the Bible from Greek and Hebrew by William Tyndale.[2]

There is a consensus that the source languages for Tyndale's translation, Biblical Greek and Hebrew, show a high frequency of CNCs (see, among others, Gianollo & Lavidas 2013, 2014; see also Blass et al. 1961). The equivalents of examples (2a–c) in the Greek New Testament (NT) show a CNC too, and in that respect, these examples demonstrate a direct translation effect. Note that not every CNC in the Greek source text has a CNC-correlate in Tyndale's NT, as can be seen from (3a) and (3b). With regard to Old English translations of Latin texts, Taylor (2008) has argued that biblical translations are different from other translations in that they may also show an indirect translation effect, whereby they follow particular characteristics of the source language even in contexts where there is no source for these characteristics. It will be seen that while Tyndale's translation demonstrates many examples of direct translation effects, CNCs behave according to the principles of "sacral stamp" (Psaltes 1913; Drinka 2011) – that is, as a biblical feature based on archaic/early English elements, which functions as a marker for the particular register. CNCs can thus be considered a biblical stylistic choice or an English biblical register marker/"sacral stamp" (see Section 3.2 on the nature of the sacral stamp and how it differs from the indirect effects of translation).[3]

---

2. Previous studies on earlier English translations have analyzed Old English with regard to the accusativus-cum-infinitivo (among others, Fischer 1994), head initial/final PPs (Taylor 2008), or the dative absolute (Timofeeva 2012). For the study of translations in historical research, see, among others, Luraghi & Cuzzolin (2007).

3. A contrast between natural language characteristics and change, on the one hand, and conscious linguistic selection, on the other, has already been shown for grammatical phenomena too (besides lexical ones). For instance, the retention and the reintroduction of affirmative *do* in 17th-century documents (and mainly in the language of liturgy) have been analyzed as a conscious selection; see, among many others, Nevalainen (1991) and Hundt (2015).

> […] given this general decline, why did *do* nevertheless sometimes survive, or even reappear, in affirmative declaratives in certain contexts after the mid-sixteen hundreds? Why did the archaic usage prevail in the language of liturgy, for instance? […] This process of more or less conscious linguistic selection is commonly contrasted with natural language change that originates from below the level of social awareness. Whether the decline of *do* in 17th-century written documents resulted from a natural change is a matter of some dispute (cf. Rissanen 1985). All we can say is that affirmative *do* was not codified as an obligatory tense and mood carrier in declaratives in standard English. Hence its retention in, or introduction into, certain contexts must have been to some extent a matter of conscious selection.   (Nevalainen 1991: 303)

(2) a.   *…but **iudge** rightewes **iudgement***
          'lit.: but judge a right judgment'   (Tyndale, John 7:24)

     b.   *…the **love** wher wt thou hast **loved** me*
          'lit.: the love with which you have loved me'   (Tyndale, John 17:26)

     c.   *…**synne** a **synne** that is not vnto deeth*
          'lit.: sins a sin that does not lead to death'   (Tyndale, 1 John 5:16)

(3) a.   *…they were **marvelously glad***   (Tyndale, Matthew 2:10)

     b.   *…**ekhárēsan kharàn** megálēn   sphódra*
          rejoiced.3PL   joy.ACC great.ACC exceedingly
          'lit.: rejoiced exceedingly with great joy'   (Greek NT, Matthew 2:10)

This paper is structured as follows. Section 2 provides the theoretical and historical background to CNCs and cognate nouns, with Section 2.1 discussing the types of cognate nouns and CNCs and Section 2.2 presenting the development of CNCs in English. Section 3 presents the analysis of the data from Tyndale's translation. The corpus study in Section 3.1 of CNCs in Tyndale and other biblical translations shows that Tyndale does not imitate the CNCs of the source language; instead, Tyndale's translation shows a preference for the archaic syntax of cognate nouns, i.e. in a PP, over the cognate direct object. In Section 3.2, I embed this archaic characteristic in Tyndale's general tendency towards archaisms in his translation (as evidenced, for instance, by his dispreference for auxiliaries). The CNCs in Tyndale's NT do not provide evidence for biblical indirect translation effects. While they are restricted to contexts where the source text has a CNC as well, they function in a way that supports their establishment as a marker of the English biblical register. Finally, Section 4 summarizes the main conclusions of the study.

2.   Theoretical and historical background

2.1   Theoretical background: Cognate nouns and CNCs

On the assumption that thematic roles can only be assigned by transitive verbs, Jones (1988) advocated the view that cognate nouns with intransitive verbs are equivalent to adverbials. Yet Massam (1990), among others, argued in favor of an argument-analysis of cognate nouns with intransitives: for instance, in some cases, passivization is available (e.g. *A song was sung by the students*), and some cognate nouns are sensitive to the number of arguments a verb may take – a property that cannot be linked to adjunct behavior. Indeed, if cognate nouns (like *fight* in (4a)) behaved like adverbials, they could co-occur with other arguments; as (4b) shows, this is not possible.

(4)   a.   They *fought* a heroic *fight*.
      b.   They *fought* the enemy heroically/ *a heroic *fight*.

Pereltsvaig (1999, 2002) and Nakajima (2006: 677) have shown that argument-like as well as adverbial cognate nouns exist: argument-like cognate nouns can be passivized, whereas adverbial cognate nouns cannot. Accordingly, they hold that cognate nouns with intransitives may appear in two possible structural positions: in an object/argumental or in an adverbial/adjunct position. This means that two alternative readings are available with cognate nouns: an adverbial/manner reading (*slept a very sound sleep* → 'slept deeply') and a non-adverbial reading (*slept a very sound sleep* → 'had a deep and sound sleep').[4]

Following up on this discussion on the different types of cognate nouns (argument-like and adverbial cognate nouns), Horrocks and Stavrou (2010) distinguished three main types of cognate nouns:

i.   cognate nouns with transitives that express referential objects and have an argumental status: these cognate nouns have concrete meanings (they can be singular or plural) and can freely be passivized (see (5));
ii.   cognate nouns with intransitive (unergative) verbs, as in (6) ("transitivizing cognate nouns"); these are also referential and have an argumental status: the passivization of cognate objects with intransitive (unergative) verbs is possible even though not common;
iii. aspectual (or activity/event-noun) cognate nouns that are non-referential; they occur with intransitive verbs and disallow passivization.[5]

The examples in (7a–b) clearly distinguish between cognate nouns of the second and the third type: while transitivizing cognate nouns (*sing a song*) can be replaced by a similar (for instance, a synonymous or hyponymous) noun (*sing the anthem* or *an aria*), aspectual cognate nouns cannot be replaced by similar nouns; their main function is the expression of a limited event with beginning and end. This function is similar to the function of light-verb constructions (for instance, *give us a grin* – instead of *grin*).

---

4.   On the question of the type of intransitive verbs, i.e. unergatives or unaccusatives, that can appear with a cognate noun, I follow Lavidas' (2013b) analysis: I claim that only unergatives can take cognate nouns of an argumental (or adverbial) character; unaccusatives only take cognate nouns of an adverbial character. In this respect, my discussion mainly concerns unergatives – but similar tendencies for expansion of CNCs hold for unaccusatives, as well.

5.   The presentation of the characteristics of the cognate nouns is brief because of space restrictions; see, among others, Jones (1988), Massam (1990), Mittwoch (1998), Pereltsvaig (1999). For a detailed analysis of the characteristics of aspectual cognate objects, see Horrocks & Stavrou (2010: 303–307).

(5)   She ended the class by asking us to *write a writing* within a given time.

(6)   They *sang* a *song* I didn't recognize.

(7)   a.   She *smiled* a (winning) *smile* (*a grin).
       vs.: b.   sang the *song/the anthem/ an aria*

Modern English has cognate nouns of the three types: cognate nouns of the first type with transitives (*write a writing/paint a painting*), as well as cognate nouns that appear in transitivizing CNCs (2nd type) and aspectual activity/event-CNCs (3rd type). See the examples of CNCs of the 2nd and 3rd type in (8a–b) (taken from Horrocks & Stavrou 2010: 293). The aspectual CNCs in Modern English of the type *smiled a winning smile, coughed a sinister cough* are not available in all languages. As will be seen in Section 2.2, aspectual CNCs with a cognate direct object were not available either before the Late Middle English period.

(8)   a.   Transitivizing CNCs
            i.   She dreamed the most wonderful dream.
            ii.   They danced all sorts of dances at that party.
            iii. Mary sang her favourite song on Peter's birthday.
       b.   Aspectual CNCs
            i.   He smiled an enigmatic smile.
            ii.   She slept a dreamless sleep.
            iii. John coughed a sinister cough.   (Horrocks & Stavrou 2010: 293)

The transitivizing CNCs comprise intransitive verbs that can behave as transitives and appear with objects that are cognate nouns or hyponyms and nouns with related meanings. In this sense, the transitivizing CNCs do not differ from constructions with transitive verbs: their cognate nouns are referential objects with a typical argumental status. In contrast, CNCs with aspectual (activity/event) cognate nouns are not similar to constructions with transitive verbs. Modern English aspectual cognate nouns are not referential nouns and they are fully selected by the verb in that they cannot be replaced by any other noun (hyponym or of related meaning). Consider (9a) vs. (9b) from Horrocks & Stavrou (2010: 294), which illustrates the difference between transitivizing and aspectual CNCs. According to Horrocks and Stavrou's (2010) analysis, the non-referential nature of the aspectual cognate nouns gives rise to the type of syntactic behavior that can be illustrated in (9b): on their view, aspectual CNCs (e.g. *grin a grin*) are similar to light verb constructions such as *give a grin*, with the only difference that aspectual CNCs provide additional information on the means by which the event takes place; the similarity resides in the fact that both light verb constructions and aspectual CNCs "effect a shift of aspectual character" with respect to the corresponding intransitive verb used on its own. In particular, in adding an object to a verb, these event – noun periphrases (light verb

constructions or constructions with aspectual cognate nouns) in Modern English turn a VP that expresses an atelic event into a telic event (in a similar fashion, non-terminative transitives (e.g. *read*) may assume a terminative interpretation when combined with a direct object (e.g. *read the book*)).[6] Horrocks & Stavrou (2010) view these periphrastic constructions that shift telicity in languages such as Modern English as a correlate of outer (viewpoint/grammatical) aspect: the telicity-shift-constructions express aspectual information in the case of absence of an outer aspect marking in the verbal stem. The aim of the present study is not to provide a further account of the correlation between aspect and telicity-shift-constructions. We may only add that outer aspect can also be expressed by means of a prefix (as in the case of Old English *ge*-; see McFadden 2015, among others), with the prefix functioning in a way similar to the verbal stem that encodes the aspectual information.

(9) a. Transitivizing CNCs
- i. *A song was sung by the revellers.*
  [+ Passivization]
- ii. *A patriotic song, everyone wanted to sing.*
  [+ Topicalization]
- iii. *What did everyone sing?*
  [+ Questioning]
- iv. *They sang (all the, some, many…) songs.*
  [+ Plural (and quantification)]
- v. *They sang a (patriotic) song.*
  Optional qualification (indefinite)
- vi. *They sang the/this/that (famous) song.*
  Optional qualification (definite)
- vii. *They sang a song, an anthem, an aria,…*
  Not restricted to cognates only

b. Aspectual CNCs
- i. *??A (winning) smile was smiled by the winner.*
  [−Passivization]
- ii. **A (winning) smile, no one smiled.*
  [−Topicalization]
- iii. **What did she smile?*
  [−Questioning]

---

6.   Terminative is not synonymous to telic for Horrocks and Stavrou: verbs that are lexically terminative in aspectual character denote a process that has a built-in end-point (e.g. terminative *melt* vs. the non-terminative *wash*). In case of a non-terminative verb, the activity may continue indefinitely (Horrocks & Stavrou 2003, 2007).

iv. *ʾShe smiled (all the, some, many…) smiles.*
[−Plural (and quantification)]

v. *ʾShe smiled a smile (cf. a winning smile).*
Qualification preferred (indefinite)

vi. *ʾShe smiled the smile (cf. the smile she was famous for).*
Qualification preferred (definite)

vii. *She smiled a (winning) smile (*a grin, *a laugh,…).*
Restricted to cognates only   (Horrocks & Stavrou 2010: 294)

To summarize, aspectual cognate nouns in Modern English are non-referring nouns which occur in CNCs expressing terminative events (in the default cases), as the terminative correlate of non-terminative intransitive verbs. On this approach, Modern English bare roots (which, for instance, do not encode the grammaticalized opposition of perfective vs. imperfective aspect in the verbal stem) may employ aspectual cognate nouns to express the preferred aspectual (terminative) character for the VP. In the following section, in which the development of CNCs in English is discussed, it will be seen that CNCs are expanded with aspectual cognate nouns, mainly from Middle to Early Modern English.

2.2   Historical background: Cognate nouns and their diachrony in English

In his historical grammar, Visser (1963–73 [2002]) discusses the development of cognate nouns with transitive and intransitive verbs. Modern English examples of such combinations of cognate nouns with transitive verbs are in (10a) and with intransitive verbs in (10b):

(10) a.   tell a tale – do a deed
b.   live a life – laugh a laugh

Visser's data (1963–73 [2002]: 415) show that in Old English, transitives are usually found with dative cognate nouns, as in (11). This early stage also saw the occurrence of CNCs including an intransitive verb and a cognate noun, but the cognate nouns took the dative case or were included in a PP.

(11) *se   Cynewulf oft   miclum   gefeohtum feaht   uuiþ*
the.NOM Cynewulf often great.DAT fights.DAT fought.3SG against
Bretwalum.   (Chron. an. 755 [Ms. A])
Britons
'Cynewulf fought many great battles against Britons.'

It is in the later stages of English, then, that Visser observed a clear tendency for an increase in the use of cognate direct objects with intransitives (instead of CNCs with cognate nouns in PPs):

> Cognate objects of the type dealt with here [with intransitives] are somewhat rare
> in Old English; they are met with increasing frequency in Middle English, and
> become quite numerous in the Modern period, where the usage, however remains
> confined to literary diction.   (Visser 1963–73 [2002]: 415)

The emergence of CNCs that include an intransitive verb and a cognate object (12;
live + Preposition + life > live life) is, for Visser, the result of analogy with transitive
verbs with cognate nouns or of the loss of the dative case.

(12)   Middle English
       *I mai **liue** a wel god **lif***   (Seven Sages, 1657)

In sum, then, cognate nouns with intransitive verbs (in contrast to cognate nouns
with regular transitives) were shown by Visser (1963–73 [2002]) to be a relatively
recent development in English: they were rare in Old English but were used with
increasing frequency in Middle and Early Modern English, and are frequent in
Modern English (see also the results from a corpus study below).

   In order to test Visser's hypotheses, I carried out a limited study of the diachrony
of CNCs in English (see Lavidas 2013a). To that effect, I collected 59 verbs that can
combine with a cognate noun; verbs were taken from the various periods of the
English language and for each verb, there was a clear description (or example) of
its first appearance in a CNC in the online version of the *Oxford English Dictionary*
(*OED*) (see Table 1).[7] The results of the corpus study are largely in line with Visser's
observations: they reveal a tendency for a more productive use of CNCs from the M2
stage (1250–1350) onwards (see the examples in (13), taken from Lavidas 2013a).[8]
This development is shown in Table 2, which presents a clear contrast between Old and
Modern English that derives from the expansion of CNCs in later stages of English.[9]

---

7.   All quantitative data are based exclusively on information in the *OED*; this means that even in a
case where Visser provides an example from a period for which the *OED* has no example that first
appeared, this period would contain no evidence of the emergence of a new CNC in the table.

8.   I follow the *York-Toronto-Helsinki Parsed Corpus of Old English Prose* (*YCOE*), the *Penn-
Helsinki Parsed Corpus of Middle English (PPCME)*, the *Penn-Helsinki Parsed Corpus of Early
Modern English (PPCEME)*, and the *Penn Parsed Corpus of Modern British English (PPCMBE)*,
with regard to chronological periods in all figures and tables: O2 (Old English; 850–950); O3 (Old
English; 950–1050); O4 (Old English; 1050–1150); M1 (Middle English; 1150–1250); M2 (Middle
English; 1250–1350); M3 (Middle English; 1350–1420); M4 (Middle English; 1420–1500); E1
(Early Modern English; 1500–1569); E2 (Early Modern English; 1570–1639); E3 (Early Modern
English; 1640–1710); MBE (Modern British English; 1700–1914).

9.   The aspectual cognate nouns belong to the newer types of cognate nouns, which expanded
the use of CNCs. A hypothesis could be that this expansion through the new aspectual cognate
nouns is related to a change in aspect (grammaticalization of the progressive) in the same stage
of the history of English. This hypothesis (namely that the grammaticalization of the progressive

(13)   a.   *The ladye **lough** a loud **laughter**, As shee sate by the king.*
       'lit.: The lady laughed a loud laughter, as she sat by the king.'

(King Estmere, l. 235 in D. Laing Early
Sc. Metrical Tales (1889), 245, c1470)

   b.   *The gaunt hobbledehoy **grinning** a very unlovely **grin**.*

(G. A. Sala, Journey due South (1887), i. xxvi. 356, 1884)

   c.   *Mr. Weller junior **smiled** a filial **smile**.*

(Dickens, Pickwick Papers (1837), xxiii 238, 1836)

Table 1. Number of CNCs that first appeared in each period (according to etymological information included in the online *OED*); from Lavidas (2013a)

Periods	O2 (OE; 850–950)	O3 (OE; 950–1050)	O4 (OE; 1050–1150)	M1 (ME; 1150–1250)	21/59			16/59			
					M2 (ME; 1250–1350)	M3 (ME; 1350–1420)	M4 (ME; 1420–1500)	E1 (EME; 1500–1569)	E2 (EME; 1570–1639)	E3 (EME; 1640–1710)	MBE (1700–1914)
New CNCs	0	0	3/59 (5.08%)	0	7/59 (11.86%)	8/59 (13.56%)	6/59 (10.17%)	1/59 (1.69%)	7/59 (11.86%)	8/59 (13.56%)	19/59 (32.20%)

may play a significant role in the emergence of the new aspectual cognate nouns) finds support in the chronology of the development of the progressive in English. The English progressive is a Late Middle English innovation, and its use increased especially during the 19th and 20th centuries (van Gelderen 2004: 203–213). Van Gelderen argues that by the time of Jane Austen's early 19th-century novels, the use of the progressive is modern-like (the progressive can be used for intervals, as in (i)). On her view, the first shift (from inner to outer aspect, with emergence of the progressive) occurs at the end of the Middle English period, whereas the shift to the fixed pattern of obligatory -*ing* occurs in the 19th century.

   (i)   *At this moment … Mr Elton **is shewing** your picture…*

(Jane Austen, *Emma I*, Chapter 7; example taken from van Gelderen 2004)

In Lavidas (2013a), I presented a corpus study with regard to the presence of forms in -*end(e/ an)* and -*ing(-yng(e)),* coded as markers for the progressive. The corpus study confirms the development of the progressive as described in the relevant studies. Moreover, it appears that the aspectual cognate nouns progress in parallel with the changes in the progressive. The stages M2 (1250–1350) and M3 (1350–1420) are the most significant stages for the development of both aspectual cognate nouns and the progressive: the new aspectual cognate nouns are productively used after stages M2 and M3, and the progressive is a Middle English innovation.

Table 2. CNCs in old and modern English

	Cognate nouns with transitives – "write a writing"	Transitivizing CNC – "sing a song"	Aspectual CNC – "grin a grin"
Old English	√	√	*
Modern English	√	√	√

In Section 2.2, I have discussed aspects of the development of the CNCs. It has been seen that Early Modern English is a significant period for the expansion of the CNCs through the aspectual cognate nouns. It is in this period (16th century) that the first English translation of the Greek and Hebrew text of the Bible, Tyndale's translation, appeared. As there is a consensus that the CNC is a characteristic construction marking the register of Biblical Greek and Hebrew, and that it is used very frequently in the Greek and Hebrew Bible texts, it will be revealing to examine CNCs in Tyndale's translation; this will be done in comparison with earlier translations (from Latin) and with later translations. While it might be expected that the CNCs in Tyndale reflect the new tendency for cognate direct objects (aspectual cognate nouns) or are only a translation effect, it will be shown that the CNCs should be analyzed as a register/biblical marker linked to archaic characteristics of English.

3.   The data from Tyndale

3.1   CNCs in Tyndale vs. other biblical translations

Tyndale's NT is the first complete English translation of the Greek NT (probably influenced by Luther's translation as well – see, among others, Light & Wallenberg 2015). The Greek NT is considered a text which is typical for the presence of CNCs, under the influence of Biblical Hebrew[10] or of biblical language in general. The Greek NT was translated by Tyndale in the first stage of the Early Modern period (1500–1569; stage E1), and Tyndale's translation was published in 1525. This means that the translation coincides with the stage of the rise of aspectual CNCs[11] (see

---

10. This influence is most clear in the Greek Old Testament / Septuagint translation from Hebrew.

11. Tyndale's Bible includes instances of the new progressive as well (see fn. 9 for the hypothetical relationship between aspectual cognate nouns and aspect); examples from Tyndale are in (ia–f). A corpus study with regard to the presence of the new progressive (-*ing* vs. the early -*ende*) demonstrates exactly the same results for Tyndale, on the one hand, and other texts of his period (E1 stage), on the other; see (ii)

Visser 1963–73 [2002] and the results from Lavidas' 2013a corpus study above). Tyndale's text influenced later English biblical translations, mainly the King James Bible. It should also be noted that Tyndale's aim was to prepare a translation that could be accessed by ordinary people; not all translators of the period shared this ambition with Tyndale.

Drawing on the source text (the Greek NT) and examples of CNCs in grammars of Biblical Greek (mainly, Conybeare & Stock 1905 [1995]; Moulton 1908; Robertson 1919; Blass et al.), I compiled a list of CNCs (for some examples, see Table 3) and then searched for all their occurrences in the electronic text of the Greek NT (GNT) (1st century AD – Koine Greek; *PROIEL Corpus*).[12] All examples thus collected were compared with their translations in Tyndale's NT (T). For comparative purposes, the following translations were examined as well: (i) the late 4th-century Latin Vulgate (LV), used for a long period in England; (ii) the Wessex/ West Saxon Gospels (WS) (circa 990 AD), that is, the earlier Old English partial translation; (iii) the Wycliffite Bible (W) (1395), that is, the earlier, Middle English, translation from the Latin text; and (iv) Luther's translation (Lu) (1522; *September Testament*), which constitutes the other possible source for Tyndale's NT. CNCs were also searched in the later English biblical translations, which were produced after Tyndale's translation: (v) the King James Version (KJV), 1611; (vi) the Young's Literal Translation (YLT), 1887; (vii) the American Standard Version (ASV), 1901; and (viii) the English Standard Version (ESV), 1971.

(i)   a.   *And he **was buyldinge** a cyte* (Tyndale, OT, Genesis 4:17)[11]

       b.   *and ther **were stondynge** theare, six waterpottes of stone after the maner of the puri-fyinge of the Iewes, **contaynynge** two or thre fyrkins a pece.* (Tyndale, John 2:6)

       c.   *Martha assone as she hearde that Iesus **was comynge**, went* (Tyndale, John 11:20)

       d.   *and no man **be dwelling** therin* (Tyndale, Acts 1:20)

       e.   *And ther **were dwelling** at Ierusalem Iewes, devoute men, which were of all nacions vnder heaven.* (Tyndale, Acts 2:5)

       f.   *And ther **were** in the same region shepherdes **abydinge** in the felde and **watching** their flocke by nyght.* (Tyndale, Luke 2:8)

(ii)   *Tyndale's NT*: 100% -*ing* vs. 0% -*ende*; but with a low frequency (presence of the new progressive: 0.28% of all clauses).

      Other E1 Texts: 100% -*ing* vs. 0% -*ende*; but with a low frequency (presence of the new progressive: 0.30% of all clauses)

12. The *PROIEL Corpus* is a parallel corpus of the old Indo-European translations of the New Testament. It was created to facilitate the study of Pragmatic Resources in Old Indo-European Languages; see Haug et al. (2008), <http://www.hf.uio.no/ifikk/english/research/projects/proiel/>, and <http://foni.uio.no:3000/users/sign_in>

Table 3. Examples of CNCs in the Greek NT for the comparative corpus study

Cognate NP in the accusative

a. *ephobéthēsan phóbon mégan*
   feared.3PL **fear**.ACC great.ACC

b. *ekhárēsan kharàn megálēn sphódra*
   rejoiced.3PL **joy**.ACC great.ACC exceedingly

c. *aúksei tền aúksēsin toû theoû*
   grow.3SG ART.ACC **growth**.ACC ART.GEN God.GEN

d. *ekaumatísthēsan hoi ánthrōpoi kaûma méga*
   burnt.3PL ART.NOM men.NOM **burn**.ACC great.ACC

e. *hamartánonta hamartían*
   sin.PTCP **sin**.ACC

f. *tò báptisma hò egồ baptízomai*
   ART.ACC **baptism**.ACC REL.ACC 1SG.NOM **baptize**.MP.1SG

g. *hē agápē hền ēgápēsás me*
   ART.NOM **love**.NOM REL. **loved**.2sg 1SG.ACC

h. *oikodomoûnti oikían*
   build.PTCP.DAT **building**.ACC

i. *euangélion aiốnion euangelísai*
   gospel.ACC eternal.ACC **evangelize**/preach.INF

j. *tò hórama hò eîden*
   ART.ACC **vision**.ACC REL.ACC **saw**.3SG

k. *tền dikaían krísin krínete*
   ART.ACC right.ACC **judgment**.ACC **judge**.2PL

l. *tês klếseōs hês eklếthēte*
   ART.GEN **calling**.GEN REL.GEN **call**.MP.2PL

m. *speîrai tòn spóron autoû*
   sow.INF ART.ACC **seed**.ACC 3SG.GEN

Cognate NP in the dative

a. *epithumíai epethúmēsa*
   desire.DAT **desired**.1SG

b. *apeilēi apeilēsómetha*
   warning.DAT **warn**.1PL

c. *parangelíai parēngeílamen*
   order.DAT **ordered**.1PL

d. *anathémati anethematísamen*
   curse.DAT **cursed**.1PL

e. *proseukhêi prosēúksato*
   prayer.DAT **prayed**.3SG

f. *têi eleutheríai hēmâs Khristòs ēleuthérōsen*
   ART.DAT **freedom**.DAT 1PL.ACC Christ.NOM **set.free**.3SG

The basis for the corpus study consists of 74 CNCs in the Greek NT. Forty appear in the accusative in the Greek source text (e.g. *ephobḗthēsan phóbon* 'feared fear. ACC'), 16 in the dative (e.g. *epithumíai epethúmēsa* 'desire.DAT desired'), and only two instances of cognate nouns occur in PPs (preposition + cognate noun in the accusative); 16 cognate nouns are attested in a relative clause (e.g. *agápēn hḕn ēgápēsen* 'love.ACC REL.ACC loved.3SG'). The majority of the Greek NT cognate nouns are modified: 62 modified vs. 12 non-modified (on the non-modified dative cognate nouns, see Gianollo & Lavidas 2014).

Tyndale's translation appears not to be slavish with regard to CNCs. It follows the Greek text only in 37.84% of the cases,[13] with *a substantial number of the cognate nouns (16/28) attested in PPs and not as cognate direct objects*. This mainly demonstrates an archaic (early English) feature of CNCs, rather than a translation effect: indeed, as pointed out above, Visser (1963–73 [2002]) has shown that in earlier stages of English, cognate nouns in CNCs with intransitive verbs were included in a PP (or marked with the dative case). Tyndale makes use of this archaic feature (cognate nouns included in a PP) instead of the Early Modern English tendency for cognate direct objects. As with all biblical translators, Tyndale is constrained by his respect of the religious text and applies an archaic register that is close to earlier features of English (see (14)). This leads to a specific register of English – which will be followed up subsequently by other authors – and not to a strict imitation of the Greek text. Moreover, all cognate nouns in Tyndale's NT are modified (even in case that the Greek source text has a non-modified cognate noun).

---

13. See, for instance, (i).

(i)　John 7:24

T:　　*but **iudge** rightewes **iudgement***

GNT:　*allà tḕn　　dikaían　krísin　　　krínete*
　　　　but　ART.ACC　right.ACC　judgment.ACC　judge.2PL

LV:　　*sed justum　judicium　　judicate*
　　　　but　right.ACC　judgment.ACC　judge.2PL

Lu:　　*sondern　richtet　eyn　recht　gericht.*
　　　　but　　judge　a　　right　judgment

WS:　　*ac demað rihtne dom.*

W:　　*but deme ye a riytful doom.*

KJV:　*but judge righteous judgment.*

YLT:　*but the righteous judgment judge*

ASV:　*but judge righteous judgment.*

ESV:　*but judge with right judgment.*

(14)   James 5:17
   T: *and he prayed **in his prayer***
   GNT: *…kaì proseukhêi prosēúksato*
            and   prayer.DAT prayed.3SG
   LV: *et   oratione   oravit*
            and speech.ABL speak.PRF.3SG
   Lu: *vnd er   betet   eyn gepet*
            and 3SG.NOM   pray.3SG a   prayer
   WS: –
   W: *and **in** preier he preiede*
   KJV: *and he prayed earnestly*
   YLT: *and **with** prayer he did pray*
   ASV: *and he prayed fervently*
   ESV: *and he prayed fervently*

CNCs in the Greek NT with non-modified cognate nouns in the dative are also translated by adverbs in Tyndale's NT (as well as in the Luther's NT);[14] see (15). The adverbs can probably express focus on the verbal predicate, similar to the cognate datives (*verum focus*; see among others, Gutzmann & Castroviejo Miró 2011 and Gianollo & Lavidas 2013).

(15)   Luke 22:15
   T: *I have **inwardly** desyred to eate this ester lambe with you*
   GNT: *epithumíai epethúmēsa toûto   tò   páskha   phageîn*
            desire.DAT desired.1SG this.ACC ART.ACC passover.ACC eat.INF
            meth' humôn
            with 2PL.GEN
   LV: *Desiderio desideravi   hoc pascha manducare vobiscum.*
            desire.DAT desire.PRF.1SG this passover eat.INF   2PL.DAT.with
   Lu: *Mich   hatt   **hertzlich** verlanget dis osterlamb mit*
            1SG.ACC have.3SG eagerly   desire.PTCP this easter.lamb with
            euch   zu essen
            2PL.DAT to eat.INF
   WS: *Of gewilnunge ic gewilnude etan mid eow þas Eastron*
   W: *With desier Y haue desirid to ete with you this pask*
   KJV: *With desire I have desired to eat this passover with you*
   YLT: *With desire I did desire to eat this passover with you*
   ASV: *With desire I have desired to eat this passover with you*
   ESV: *I have earnestly desired to eat this Passover with you*

---

14. This may reflect a correlation between Tyndale's and Luther's NT. I leave the issue open for future research.

CNCs comprising cognate nouns with an indefinite article (16) and relative clause CNCs (17) are also attested frequently in Tyndale's translation.

(16)   1 John 5:16
    T: *Yf eny man se his brother **synne a synne** that is not vnto deeth*

    GNT: *eán tis   ídēi   tòn   adelphòn   autoû*
            if   someone.NOM see.3SG ART.ACC brother.ACC 3SG.GEN
            hamartánonta hamartían mề   pròs thánaton
            sin.PTCP.ACC sin.ACC   NEG to   death

    LV: *Qui   scit   fratrem   suum   peccare peccatum*
           somebody know.3SG brother.ACC 3SG.ACC sin.INF sin.ACC
           non ad mortem petet
           NEG to death   aim.FUT.3SG

    Lu: *So   yemand   sihet   seynen bruder   sundigen eyne sunde,*
           when somebody see.3SG his   brother sin   a   sin
           nicht zum todt
           NEG to   death

    WS: –

    W: *that his brother synneth a synne not to deth*

    KJV: *If any man see his brother sin a sin which is not unto death*

    YLT: *If any one may see his brother sinning a sin not unto death*

    ASV: *If any man see his brother sinning a sin not unto death*

    ESV: *If anyone sees his brother committing a sin not leading to death*

(17)   Ephesians 2:4
    T: *[But God which is rich in mercy] thorow his greate **love wherwith he loved** vs*

    GNT: *…dià   tền   pollền   agápēn   autoû   hền*
           through ART.ACC great.ACC love.ACC 3SG.GEN REL.ACC
           ēgápēsen   hēmâs
           loved.3SG 1PL.ACC

    LV: *propter nimiam   caritatem suam,   qua*
           near   excessive.ACC love.ACC 3SG.ACC REL.NOM
           dilexit   nos
           love.PRF.3SG 1PL.ACC

    Lu: *durch   seyne grosse liebe, da   mit er   vnns*
           through his   great love there with 3SG.NOM 1PL.ACC
           geliebt   hatt
           love.PTCP have.3SG

    WS: –

    W: *for his ful myche charite in which he louyde vs*

    KJV: *for his great love wherwith he loved us*

    YLT: *because of His great love with which He loved us*

    ASV: *for his great love wherwith he loved us*

    ESV: *because of the great love with which he loved us*

Table 4. CNCs in the Greek NT and its translations

Text	Greek NT	Tyndale's NT	Latin Vulgate	Luther's September Testament	West Saxon Gospels	Wycliffe	KJV	YLT	ASV	ESV
Type of CNC										
Total number of CNCs	74	28/74 [37.84%]	52/74 [70.27%]	32/74 [43.24%]	16/34* [47.06%]	38/74 [51.35%]	32/74 [43.24%]	44/74 [59.46%]	38/74 [51.35%]	28/74 [37.84%]
Accusative/ Objective (Cognate direct object)	40 [54.05%]	6 [21.43%]	10 [19.23%]	8 [25%]	4 [25%]	8 [21.05%]	6 [18.75%]	10 [22.73%]	8 [21.05%]	2 [7.14%]
Dative	16 [21.62%]	–	28 [53.85%]	0 [0%]	0 [0%]	0 [0%]	–	–	–	–
PP	2 [2.70%]	16 [57.14%]	0 [0%]	12 [37.5%]	6 [37.5%]	22 [57.89%]	14 [43.75%]	16 [36.36%]	18 [47.37%]	18 [64.29%]
Relative clause	16 [21.62%]	6 [21.43%]	14 [26.92%]	12 [37.5%]	6 [37.5%]	8 [21.05%]	12 [37.5%]	18 [40.91%]	12 [31.58%]	8 [28.57%]
Modified	62	28	40	32	12	30	30	36	36	26
Non-modified	12	0	12	0	4	8	2	8	2	2

* It contains only the Gospels.

Table 4 presents the results of the corpus study on biblical translations. We should notice that biblical translations can provide a natural parallel corpus with homogeneous pragmatic contexts and syntactic faithfulness. However, the faithfulness, observed in biblical translations, can reach an extreme degree and probably override native characteristics of the target language (Gianollo 2011). Table 4 demonstrates that the Latin Vulgate is the closest translation to the Greek source text with regard to CNCs: 70.27% of the CNCs in the Greek text have a CNC in the Latin translation as well. It also includes a large number of (non-modified) datives.

Luther's translation shows results similar to Tyndale's NT (similar percentages of CNCs; Lu: 32/74 – T: 28/74). However, it is not the case that CNCs in Tyndale's NT always correspond to the CNCs in Luther's NT; see, for instance (18), where a translation involving a cognate noun in a PP in Tyndale's translation has no cognate noun correlate in Luther's translation.

(18)   T: PP with cognate noun
            *and encreaseth **with the increasynge** that commeth of god*
      Lu: No cognate noun
            vnd also wechst   zur grosse
            and so   grow.3sɢ to   growth

All English biblical translations that appeared before Tyndale's translation were produced based on Jerome's Latin Vulgate. The West Saxon Gospels constitute a partial, Old English translation of the Latin NT: this text includes more CNCs comprising cognate nouns in a PP than cognate direct objects. PPs with cognate nouns are also typical of the earlier Wycliffite translation (from Latin).[15] Even though the Latin source mainly includes cognate accusatives and datives, many cognate nouns in the Wycliffite translation are also attested in PPs. With regard to the translations of the 19th and 20th century, YLT includes the largest number of CNCs in the English translations examined (44 out of the 74 CNCs of the Greek NT); here as well, most of the cognate nouns occur in PPs. In addition, the ESV follows Tyndale's strategy of avoiding cognate direct objects: the translators prefer cognate nouns in PPs to cognate direct objects.

On the whole, the English biblical translations, with the exception of ESV, show a fairly stable proportion of CNCs comprising cognate direct objects. Importantly, the expansion of CNCs with cognate direct objects that can be observed in the language in general is not seen in the English biblical translations (see Table 5). On the contrary, cognate nouns in PPs are preferred both before and after the loss of the dative case.

---

15. I thank the anonymous reviewer for bringing the issue of the earlier vs. later version of the Wycliffite Bible. I would like to point out that the results are based on a detailed corpus study in the earlier version. However, it appears that, in most of the cases, the later version does not modify the earlier version with regard to the CNCs.

Table 5. Cognate direct objects in English biblical translations

Text	Greek NT	Tyndale's NT	Latin Vulgate	Luther's September Testament	West Saxon Gospels	Wycliffe	KJV	YLT	ASV	ESV
Type of CNC										
Accusative/ Objective (Cognate direct object)	40/74 [54.05%]	6/28 [21.43%]	10/52 [19.23%]	8/32 [25%]	4/16 [25%]	8/38 [21.05%]	6/32 [18.75%]	10/44 [22.73%]	8/38 [21.05%]	2/28 [7.14%]

It is my claim that the tendency for a preference of CNCs with cognate nouns in PPs reflects a type of continuation of an archaic tradition (in the sense of a preference in early English for cognate nouns in PPs over cognate direct objects). The translator, instead of giving preference to the source text option (a cognate direct object / cognate accusative), which is also an available and productive option for the native language of the period, follows an option that was common in earlier English, namely a cognate noun in a PP. This preference is furthermore in contrast to the fact that cognate direct objects (cognate nouns in the accusative) are a typical biblical feature of Greek. In this respect, English also follows, in a way, the Gothic tradition; see, for instance, Wolfe (2006), who argues in favor of a tendency in Gothic to avoid cognate direct objects (accusative cognate nouns), as in (19).

(19) GNT *kaì* **ekséstēsan** **ekstásei** *megálēi*
      and astonished.3PL astonishment.DAT great.DAT
      Gothic *jah* *usgeisnodedun* *faurhtein* *mikilai*
      and astonished.3PL fear.DAT great.DAT
      'and they were astonished with great astonishment' (Mark 5:42)

The characteristics of CNCs, as they appeared in Tyndale's and later translations, can be subsumed under the notion of "sacral stamp": according to Psaltes (1913), "sacral stamp" covers the linguistic characteristics that reflect membership in the Christian community ("a linguistic emblem of membership in the Christian community" (Doron 2011: 41)). These characteristics imitate Koine Greek characteristics of the Bible and create a text that attempts to distance itself from the language of its period. In this paper, a broader definition will be used, as proposed by Drinka (2011: 67):

> Translators have shaped their style and syntax according to these patterns [the ones attested in the earlier biblical texts], as they endowed their own words with the archaic sounds of the past that connoted reverence and membership in the Christian community.[16]

Drinka has shown the role of the "sacral stamp" even in the Gospel of Luke and the Acts of the Apostles (written in Greek): Luke, for instance, adopted the archaic style of the Septuagint (the Greek translation of the Old Testament) and provided his text with what Wifstrand (2005: 42) describes as: "an aura of sacred history, making it appear as the sequel and fulfillment of the Old Testament". Another structure that has been considered an example of the "sacral stamp" is the periphrastic progressive; this structure is very frequently used in the Greek New Testament, following the Septuagint model, and then in the Latin Vulgate and the Latin Christian texts. In Amenta (2003), for instance, it is argued that the periphrastic progressive has been a structure which functions as a marker of membership in the Christian community.

On the basis of the data from Tyndale's translation and other translations of the Bible, I argue that the "sacral stamp," as a marker of adherence to the register of Christian texts, can deviate from the Greek source text to follow English rules of an earlier period (cognate nouns in PPs). The use of CNCs with cognate nouns in PPs – in contrast to the preference of the period for cognate direct objects and the choice in the source text of cognate accusatives (or datives) – constitutes the primary reason for considering these to be archaic patterns (in opposition, for instance, to the tendency in Latin to use cognate direct objects). Recall that cognate direct objects were not productive in Old English, and, in contrast, the only option in Old English was the use of a cognate noun in a PP (or in the dative case).

Section 3.2 presents additional manifestations of Tyndale's preference for archaisms. It is argued that CNCs, in combination with these other manifestations, establish an English biblical register rather than reflect a translation effect.

3.2 Archaic "sacral stamp" vs. indirect effects of translation in Tyndale

Besides direct effects of translation, i.e. the repetition/imitation of a pattern of the source text in the target text within the same context, previous studies have recognized indirect effects of translation. According to Taylor (2008), a translator may

---

16. The following remark by Drinka (2011: 47, fn.9 – my emphasis) on the role of archaic elements as aspects of a "sacral stamp" is significant for our discussion:

> Note that English, too, partakes of the "sacral stamp", in elevating these familiar lines by means of *archaic language*, such as that found in the King James version: 'My soul doth magnify the Lord/ And my spirit hath rejoiced in God my Saviour.'

prefer to use syntactic structures of the source text in the target text (in a higher frequency than other authors of the same period) even in contexts where there is no matching structure in the source text. This indirect effect of translation is a kind of priming, in that the higher frequency of particular structures may be attributed to structures that have been used in the previous discourse. Taylor has related the different types of translation effect to the type of the translation: translations of non-biblical texts show only direct translation effects; translations of biblical texts show *both* direct and indirect effects. In her analysis, this means that different translation strategies are put into practice in biblical (seen by the translator as containing the language of God) vs. non-biblical texts. As Taylor (2008: 355–356) points out, "OE writers generally felt free to paraphrase rather than give a literal translation of texts, except when it came to the Bible". For instance, following the head-initial order rule of Latin (the source language), Old English biblical translations show a higher rate of head-initial orders in PPs in general (and not only in the contexts where the source text has a PP); see Figure 1.
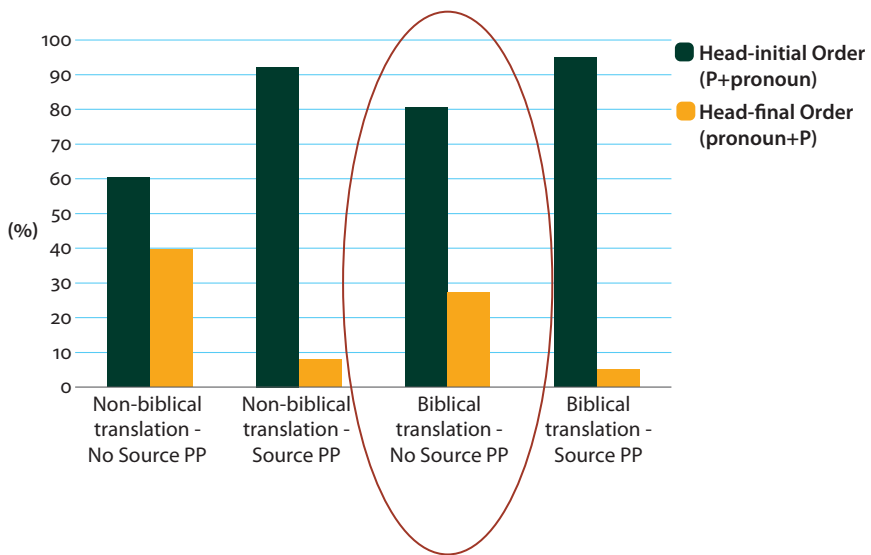
Figure 1. Old English: Frequency of head-initial order with and without a Latin source PP (Taylor 2008)

Importantly, Tyndale's NT does not constitute a typical example of a biblical translation with indirect translation effects: there are only very few cases of use of a CNC in a context where the source text does not include a CNC (see, for instance, (20)). A corpus search for CNCs in Tyndale's translation in contexts other than those of the source text yielded only very few instances and, therefore, cannot support an indirect effect with regard to CNCs.

(20)   1 Tim. 2:2
    a.   Tyndale: CNC
       *that we maye **live** a quyet and a peasable **life** in all godlines and honestie.*
    b.   Greek NT: No CNC
       *hína   éremon   kaì hēsúkhion **bíon**   **diágōmen**   en pásēi*
       COMP calm.ACC and quiet.ACC life.ACC pass.SBJV.1PL in all.DAT
       eusebeíai   kaì semnótēti.
       reverence.DAT and dignity.DAT

It appears that the frequent use of CNCs (with cognate nouns in the accusative) in the Greek source text does not force Tyndale to use CNCs productively in the target text, and especially in the case of their absence from the source text. In other words, Tyndale's NT does not show indirect translation effects (in a similar way to the biblical translations as presented, for instance, in Taylor's study) and follows non-biblical translations in this respect. CNCs in Tyndale's translations appear to function as a "sacral stamp" based on an archaic (early English) characteristic (cognate nouns in PPs and not cognate direct objects). For this reason, they do not follow the new tendency of CNCs with (aspectual) cognate direct objects in all cases, they do not demonstrate an indirect effect of translation and neither are they translated in a slavish manner.

This type of "sacral stamp" to CNCs can be embedded in Tyndale's general tendency towards archaisms in his translation: in addition to the use of CNCs, a contrast between features of the natural development of English and Tyndale's NT can be observed with regard to other grammatical phenomena as well. In Early Modern English (16th century), periphrastic modal auxiliaries (e.g. *he **should hear** them…*) replace the early inflected subjunctive. However, as can be seen from Figure 2 (from Canon 2010, based on Harsh 1968), Tyndale still frequently uses the inflected subjunctive in his works, thus retaining an archaic feature of the language. Another archaic characteristic of the language in Tyndale's texts is Tyndale's preference for the modal preterite over the new modal auxiliary.[17]

---

17. Canon (2010: 75–76) states:

> There is a profound difference between the two texts [Tyndale vs. KJV] in the categories of modal preterite and modal auxiliary, with Tyndale showing a heavier usage of the former and the KJV the latter. Percentages of unmarked tokens were the same in both, and inflected subjunctive forms were slightly more prevalent in the KJV than in Tyndale. Since the KJV closely follows the Tyndale Bible chronologically, and is largely built upon it, the shift from modal preterite usage to modal auxiliary usage is significant.
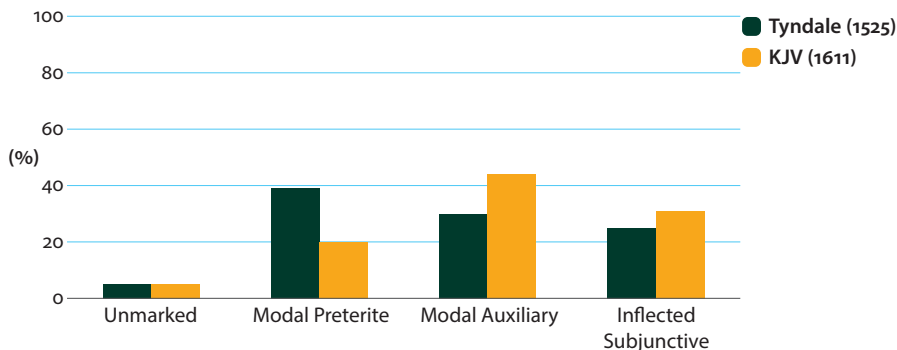
Figure 2. Modal use by type (from Canon 2010: 75, based on Harsh 1968)

Harsh's (1968) account of the differences between Tyndale's Bible and the other translations under examination emphasizes the fact that Tyndale's preference for the modal preterite over the new modal auxiliary shows an archaic element that is chosen in the language of Tyndale:[18] Tyndale demonstrates a higher percentage of modal preterites than any of the biblical translations, even the ones of the earlier stages.[19]

Summing up, it appears that cognate nouns in PPs are part of a general tendency for Tyndale to use archaisms in his translation, as evidenced, for instance, in his dispreference of modal auxiliaries. CNCs function together with other archaic elements to establish an English biblical register rather than reflect a translation effect.

4.   Conclusions

Tyndale's translation is produced during a period of expansion of CNCs through the type of aspectual cognate direct objects in English. Furthermore, the source texts of Tyndale's translation (Greek and Hebrew Bible) contain a significant number of CNCs. Accordingly, a study of Tyndale's text – in comparison with the other English

---

18. Harsh (1968) examined the Gospel translations from the Rushworth Latin, Rushworth Gloss, Wycliffe, Tyndale, King James, and Goodspeed Bibles.

19. Tyndale's translations have not yet been the focus of many linguistic studies. Therefore, there are only very few studies that refer to the archaic elements in Tyndale's text. One is, for instance, the study by Yamakawa (1970), who examines the presence of the type *anhungred/anhongred/ anhoungered/an hungred* 'hungered' in Tyndale's and other biblical translations. These forms with the archaic prefix (*an-*) appear first (with regard to English Bibles) in Tyndale's translation and reappear in subsequent English translations of the Bible (of the 16th, 17th, and 19th century). Further examination of the archaic constructions in Tyndale's translation remains open for future research.

biblical translations – can provide evidence as to the status of CNCs in Early Modern English as well as to their role in the first English biblical translation from Greek.

It has been argued that Tyndale's NT is a biblical translation that does not show indirect translation effects: a frequent use of CNCs in the translation is not observed when they are absent in the source text. Rather, CNCs in Tyndale's translation display an archaic (early English) characteristic: they appear with cognate nouns in PPs rather than with cognate direct objects. In this manner, the CNCs constitute markers of the English biblical register, a type of a "sacral stamp", which is a tendency to adhere to Christian and mainly Koine Greek linguistic features. It appears, however, that the "sacral stamp" can deviate from the source (Greek) text in following earlier English rules. It has also been pointed out that the archaic characteristic of cognate nouns in PPs is part of a general tendency for Tyndale to use archaisms in his translation, as evidenced, for instance, in his dispreference of modal auxiliaries. As such, the CNCs in Tyndale's NT function together with other archaic elements to produce an English biblical register.

Acknowledgements

Abbreviations

1/2/3	first/second/third person	LV	Latin Vulgate
ABL	ablative	MP	mediopassive
ACC	accusative	NEG	negative
ART	article	NOM	nominative
ASV	American Standard Version	NT	New Testament
COMP	complementizer	PL	plural
DAT	dative	PRF	perfect
ESV	English Standard Version	PTCP	participle
FUT	future	REL	relative
GEN	genitive	SG	singular
GNT	Greek New Testament	SBJV	subjunctive
INF	infinitive	T	Tyndale's New Testament
KJV	King James Version	WS	Wessex/West Saxon Gospels
Lu	Luther's translation (September Testament)	W	Wycliffite Bible
		YLT	Young's Literal Translation

References

Blass, Friedrich, Debrunner, Albert & Funk, Robert W. 1961. *A Greek Grammar of the New Testament and Other Early Christian Literature*. Chicago IL: The University of Chicago Press.

Canon, Elizabeth Bell. 2010. *The Use of Modal Expression Preference as a Marker of Style and Attribution: The Case of William Tyndale and the 1533 English Enchiridion Militis Christiani*. Bern: Peter Lang.

Conybeare, Frederick Cornwallis & Stock, S. George. 1905[1995]. *Grammar of Septuagint Greek*. Boston MA: Hendrickson.

Drinka, Bridget. 2011. The sacral stamp of Greek: Periphrastic constrictions in New Testament translations of Latin, Gothic, and Old Church Slavonic. In *Indo-European Syntax and Pragmatics: Contrastive Approaches*, Eirik Welo (ed.). *Oslo Studies in Language* 3(3): 41–73.

Fischer, Olga. 1994. The development of quasi-auxiliaries in English and changes in word order. *Neophilologus* 78: 137–164. doi:10.1007/BF00999959

Gianollo, Chiara. 2011. Native syntax and translation effects. Adnominal arguments in the Greek and Latin New Testament. In *Indo-European Syntax and Pragmatics: Contrastive Approaches*, Eirik Welo (ed.). *Oslo Studies in Language* 3(3): 75–101.

Gianollo, Chiara & Lavidas, Nikolaos. 2013. Cognate adverbials and case in the history of Greek. *Studies in Greek Linguistics* 33: 61–75.

Gianollo, Chiara & Nikolaos Lavidas. 2014. Greek cognate datives: From modification to focus. In *Selected papers from the 11th International Conference on Greek Linguistics*, Georgios Kotzoglou, Kalomira Nikolou, Eleni Karantzola, Katerina Frantzi, Ioannis Galantomos, Marianthi Georgalidou, Vasilla Kourti-Kazoullis, Chrysoula Papadopoulou & Evagelia Vlachou (eds), 488–500. Rhodes: University of the Aegean, Laboratory of Linguistics of the SE Mediterranean.

Gutzmann, Daniel & Castroviejo Miró, Elena. 2011. The dimensions of verum. In *Empirical Issues in Syntax and Semantics* 8, Olivier Bonami & Patricia Cabredo Hofherr (eds), 143–165. Paris: Université Paris-Sorbonne. <http://www.cssp.cnrs.fr/eiss8>

Harsh, Wayne. 1968. *The Subjunctive in English*. Tuscaloosa AL: University of Alabama Press.

Haug, Dag T. T. & Jøhndal, Marius L. 2008. Creating a parallel treebank of the Old Indo-European Bible translations. In *Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data* (LaTeCH 2008), Caroline Sporleder & Kiril Ribarov (eds), 27–34. Marrakech: European Language Resources Association.

Horrocks, Geoffrey & Stavrou, Melita. 2003. Actions and their results in Greek and English: The complementarity of morphologically encoded (viewpoint) aspect and syntactic resultative predication. *Journal of Semantics* 20: 297–327. doi:10.1093/jos/20.3.297

Horrocks, Geoffrey & Stavrou, Melita. 2007. Grammaticalized aspect and spatio-temporal culmination. *Lingua* 117: 605–644. doi:10.1016/j.lingua.2006.03.007

Horrocks, Geoffrey & Stavrou, Melita. 2010. Morphological aspect and the distribution of cognate objects across languages. In *Lexical Semantics, Syntax, and Event Structure*, Malka Rappaport Hovav, Edit Doron & Ivy Sichel (eds), 284–308. Oxford: OUP. doi:10.1093/acprof:oso/9780199544325.003.0014

Hundt, Marianne. 2015. Heterogeneity vs. homogeneity. In *Letter Writing and Language Change*, Anita Auer, Daniel Schreier & Richard J. Watts (eds), 72–100. Cambridge: CUP. doi:10.1017/CBO9781139088275.006

Jones, Michael Allan. 1988. Cognate objects and the case-filter. *Journal of Linguistics* 24: 89–110. doi:10.1017/S0022226700011579

Lavidas, Nikolaos. 2013a. Null and cognate objects and changes in (in)transitivity: Evidence from the history of English. *Acta Linguistica Hungarica* 60(1): 69–106. doi:10.1556/ALing.60.2013.1.2

Lavidas, Nikolaos. 2013b. Unaccusativity and the diachrony of null and cognate objects in Greek. In *Argument Structure in Flux: The Naples-Capri Papers* [Studies in Language Companion Series 131], Elly van Gelderen, Jóhanna Barðdal & Michela Cennamo (eds), 307–342. Amsterdam: John Benjamins. doi:10.1075/slcs.131.12lav

Light, Caitlin & Wallenberg, Joel. 2015. The expression of impersonals in Middle English. *English Language and Linguistics* 19(2): 227–245. doi:10.1017/S1360674315000076

Luraghi, Silvia & Cuzzolin, Pierluigi. 2007. Mediating culture through language: Contact induced phenomena in the early translations of the Gospels. In *Europe and the Mediterranean as Linguistic Areas: Convergencies from a Historical and Typological Perspective* [Studies in Language Companion Series 88], Paolo Ramat & Elisa Roma (eds), 133–158. Amsterdam: John Benjamins. doi:10.1075/slcs.88.08lur

Massam, Diane. 1990. Cognate objects as thematic objects. *Canadian Journal of Linguistics* 35: 161–190.

McFadden, Thomas. 2015. Preverbal *ge-* in Old and Middle English. In *Byproducts and Side Effects* [ZAS Papers in Linguistics 58], André Meinunger (ed.), 15–48. Berlin: Zentrum für Allgemeine Sprachwissenschaft (ZAS).

Mittwoch, Anita. 1998. Cognate objects as reflections of Davidsonian event arguments. In *Events and Grammar*, Susan Rothstein (ed.), 309–332. Dordrecht: Kluwer.
doi:10.1007/978-94-011-3969-4_13

Moulton, James Hope. 1908. *A Grammar of New Testament Greek, Vol. 1: Prolegomena*. Edinburgh: T. and T. Clark.

Nakajima, Heizo. 2006. Adverbial cognate objects. *Linguistic Inquiry* 37(4): 674–684.
doi:10.1162/ling.2006.37.4.674

Nevalainen, Terttu. 1991. Motivated archaism: The use of affirmative periphrastic *do* in Early Modern English liturgical prose. In *Historical English Syntax*, Dieter Kastovsky (ed.), 303–320. Berlin: Mouton de Gruyter. doi:10.1515/9783110863314.303

Pereltsvaig, Asya. 1999. Two classes of cognate objects. In *The Proceedings of the WCCFL XVII*, Kimary Shahin, Susan Blake & Eun-Sook Kim (eds.), 537–551. Stanford CA: CSLI.

Pereltsvaig, Asya. 2002. Cognate objects in Modern and Biblical Hebrew. In *Themes and Issues in Arabic and Hebrew*, Jamal Ouhalla & Ur Shlonsky (eds), 107–136. Dordrecht: Kluwer.

Psaltes, Stamatios B. 1913. *Grammatik der byzantinischen Chroniken*. Göttingen: Vandenhoeck & Ruprecht.

Robertson, Archibald Thomas. 1919. *A Grammar of the Greek New Testament in the Light of Historical Research*, 3rd edn. New York NY: Hodder and Stoughton.

Taylor, Ann. 2008. Contact effects of translation: Distinguishing two kinds of influence in Old English. *Language Variation and Change* 20(2): 341–365. doi:10.1017/S0954394508000100

Timofeeva, Olga. 2012. Latin absolute constructions and their Old English equivalents: Interfaces between form and information structure. In *Information Structure and Syntactic Change in the History of English*, Anneli Meurman-Solin, María José López-Couso & Bettelou Los (eds), 228–242. New York NY: OUP. doi:10.1093/acprof:oso/9780199860210.003.0010

van Gelderen, Elly. 2004. *Grammaticalization as Economy* [Linguistik Aktuell/Linguistics Today 71]. Amsterdam: John Benjamins. doi:10.1075/la.71

Visser, Fredericus Theodorus. 1963–73[2002]. *An Historical Syntax of the English Language*. Leiden: E. J. Brill.

Wolfe, Brendan. 2006. Figurae etymologicae in Gothic. *Oxford University Working Papers in Linguistics, Philology and Phonetics* 11: 207–214.

Yamakawa, Kikuo. 1970. 'An hundred' and some kindred syntactic archaisms. *Hitotsubashi Journal of Arts and Sciences* 11(1): 53–62.

CHAPTER 3

On the differential evolution of simple and complex object constructions in English

Günter Rohdenburg

University of Paderborn

This article surveys the evolution of object structures in Modern English. In the area of simple constructions, a vast range of prepositional objects and adjuncts have been replaced by direct objects, thus making the latter category considerably more abstract. By contrast, in the area of more complex structures, English has experienced a series of dramatic changes, leading to the virtual loss of several types of construction and the contraction of many others. Most of these reductive changes have introduced a high degree of functional specialization, by narrowing the semantic spectrum of the original syntactic frame or by compelling the use of alternative grammatical devices. The paper provides a corpus-based analysis of three major domains, external possessor constructions, double objects, and clausal complements preceded by direct objects.

Keywords: simple versus complex object structures, direct objects, prepositional objects, external versus internal possessor constructions, double objects, clausal complements, rise of constructions, loss of constructions, functional diversity, functional specialization

1.   Introduction

In his *Comparative Typology of English and German*, Hawkins (1986) shows that – compared with German – several major grammatical categories of English have become enlarged and more abstract. Inspired by Sapir (1921), Hawkins argues throughout the book in support of the following generalization:

> The drift in the history of English has clearly been towards the more extensive use of more limited formal means, with the resulting complexity in the mapping between form and meaning.   (Hawkins 1986: 129)

One of the most striking differences between English and German concerns the subject category, which had already been explored in some depth (see, e.g., Rohdenburg

© 2018 John Benjamins Publishing Company

1974). The class of inanimate and non-agentive subjects in English is very much larger than that of German (and most other European languages). Consider, for instance, the contrasts in (1)–(3).

(1) a.   So, that mows the lawn. (What shall I do next?)
     b.   *So, das mäht den Rasen.*

(2) a.   This material tailors 5 dresses.
     b.   *Dieser Stoff schneidert 5 Kleider.*

(3) a.   The book had sold over 100,000 copies.
     b.   *Das Buch hatte über 100,000 Exemplare verkauft.*

To date, there are no equally detailed investigations concerning the range of object constructions.[1] Even so, in common with other linguists (e.g. Plank 1983: 9–11, 1984: 342ff; König & Gast 2007: 106–107), Hawkins (1986: passim) assumes that the domain of object constructions has witnessed a similar evolution to that of the subject category. The purpose of this paper is to show that, depending on the kind of object construction in question, English has evolved in two diametrically opposed ways. In the area of simple constructions featuring just one zero-coded nominal complement, the direct object has indeed undergone a similar expansion to the subject category. Thus, as in (4), numerous prepositional phrases have been largely replaced by direct objects.

(4) They had joined forces to protest (against) the poll tax.

As a result, the functional diversity of the direct object has increased immensely. However, in the area of more complex object constructions, involving the co-presence of two zero-coded nominal complements or a direct object followed by a clausal complement, we can observe centuries-old and strikingly contrary trends. For instance, double objects after the verb *deliver* are no longer acceptable today in cases like (5), where agentive subjects co-occur with concrete objects. Furthermore, with several verb classes, combinations of direct object and *that*-clause as in (6) have also been phased out.

(5) I deliver'd him his Cargo, …   (*Eighteenth Century Fiction* (*ECF*), 1723)

(6) Saladin answered him, that he also ruled over as many peoples.
                                              (*Oxford English Dictionary* (*OED*), 1639)

---

1.   Previous research dealing with the tendency to convert intransitive verbs into transitive ones includes the following: Jespersen (1927: 252–273), Kirchner (1955, 1957, 1959), Rohdenburg (1974: 79–83, 357–411, 1995a: 105–108, 2009a: 197–201), Legenhausen (1988).

The resulting changes have usually led to a reduction and functional specialization of various argument complexes. The present study also includes a number of examples instantiating the frame S – V – O – PP, where the prepositional phrase represents a construction-specific element. Owing to divergent internal features, these constructions have been assigned to either the area of simple constructions in Section 2 or to that of the complex ones in Section 3. Arguments motivating the classification adopted will be supplied in Section 3. The analyses presented in this paper are generally confined to British English ranging from the Early Modern period to the present day. The database used for this study consists of the *British National Corpus* (*BNC*), *OED2*, a few years of *The Times* and *The Sunday Times* as well as a sizeable collection of historical British datasets provided by Chadwyck-Healey and the Gutenberg Project (see the bibliography for further details).

2. Simple object structures: Increased functional diversity

This section deals with the massive changeover leading from prepositional objects or adjuncts to direct objects. The area is so vast that we can only touch upon three aspects – out of a total of well over a dozen.

2.1 Directive verbs

In earlier stages of the language, a large number of directive (or manipulative) verbs used to select directional prepositions in the two environments shown in (7), (8), and (9).

(7)   a.   He incited (to) a rebellion.
      b.   He incited them to a rebellion.

(8)   a.   The doctor advised (to) an operation.
      b.   The doctor advised her to an operation.

(9)   a.   They urged (to) an immediate withdrawal.
      b.   They urged us to an immediate withdrawal.

Here, we are dealing with prepositional phrases functioning as effected rather than affected patients. In the simpler cases illustrated by (7a), (8a), and (9a), the prepositional phrase has been generally replaced by a direct object, which also represents an effected patient. Some of the directive verbs displaying this development in former times include those listed in (10).

(10)   *advise, compel, counsel, encourage, enjoin, entreat, incite induce, inspire, invite, solicit, stimulate, urge*

However, to the extent that the more complex cases as in (7b), (8b), and (9b), featuring an additional – and typically human – experiencer, have survived at all, they have invariably retained the prepositional phrase. As a result, the structural gap between the simpler and more complex structures has been widened. By contrast, in languages like German, the prepositional equivalents in both the a- and b-examples have been preserved throughout. Table 1 sketches the evolution of the three types of construction with the verb *incite*, where the decline of the prepositional object in cases like (7a) can be traced as far back as Early Modern English.

Table 1. Direct and prepositional objects after *incite* in a series of historical databases[2]

		I V+*to* NP	II V+direct object	III V+NP1 *to* NP2	IV total
1	16th–17th c. (EEPF, EPD)	3 (8.8%)	8 (23.5%)	23 (67.6%)	34
2	18th c. (ECF, EPD, NCF)	4 (9.8%)	15 (36.6%)	22 (53.7%)	41
3	19th c. (NCF, EPD, MNC/B, LNC/B)	10 (16.4%)	11 (18.0%)	40 (65.6%)	61
4[*]	1960s–1993 (BNC: fictional domain [= wridom1])	1 (5.9%) (legal context)	13 (76.5%)	3 (2/1) (17.6%)	17
5[*]	1960s–1993 (BNC: overall)	5 (4.2%) (legal context)	98 (81.7%)	17 (13/4) (14.2%)	120

* The bracketed figures in rows 4 and 5 distinguish between *to* and *into,* a minority variant attested only in Present-day English.

The presentation of the data starts at a stage where the direct object in column II dominates already over the – presumably older – *to*-phrase in column I. From then on, the overall situation remains relatively stable for at least the next two centuries. The 20th century must have witnessed some important changes. To begin with, there are striking developments involving the more complex construction of type (7b) in column III and the direct object of type (7a) in column II, with the proportion of the more complex structure declining dramatically and that of the direct object increasing in a corresponding manner. Furthermore, since all of the remaining prepositional objects in the simple case (column I) are restricted to formal legal contexts, we can assume that, just as with *advise* and *urge* in examples like (8a) and (9a), the *to*-phrase has practically disappeared from ordinary language.

In English – and again unlike languages like German – many similar divergences can be found in several subdomains, featuring two major structural choices: (i) the direct object, which in some cases (still) alternates with a prepositional object or adjunct, and (ii) the combination of direct object + prepositional object or adjunct. In illustration, consider:

---

2. Full references to the electronic sources in the tables can be found at the end of the paper, following the Reference section.

(11) a.   They threatened (*with) violence.
      b.   They threatened him with violence.

(12) a.   She won (on) the pools.
      b.   She won a lot of money on the pools.

(13) a.   This entitles (to) a refund.
      b.   This entitles you to a refund.

In all three cases, it can be assumed that the establishment of the direct object in the a-examples has succeeded a prepositional stage. While in example (11a) the direct object has been around for some time, those in (12a) and (13a) have only been introduced in more recent times. Unlike (12a), the direct object in (13a) is as yet hardly attested in the language of British newspapers.

2.2   Body part instruments

Next we turn to the vast set of verbs which involve agentives activating parts of their body in characteristic ways as, for instance, in (14).

(14)   She clicked her tongue.

In the literature, object NPs denoting such moving entities have usually been referred to as instrumental objects. The reason is immediately apparent. In many European languages including Dutch and German, the translation equivalents of examples like (14) tend to involve – to this day – direct equivalents of *with*-phrases or instrumental case forms (Bibovič 1976). In English, by contrast, nearly all such body part instruments are usually realized today as direct objects. Crucially, if we go back in time, we find that instrumental *with*-phrases have been available with all of the verbs in (15) either sporadically and predominantly or even exclusively.

(15)   *beat-wing, blink-eye, clap-hand, clack-tongue, click-tongue, cluck-tongue, flap-wing, gnash-tooth, grind-tooth, kick-leg, loll-shoulder, nod-head, pout-lip,* **point – finger**, *shrug-shoulder, shuffle-foot, snap-finger,* **stamp – foot, tap – finger/foot. wave – hand/ arm. wink – eye**, etc.

Two of such older uses that have become obsolete today are presented in (16) and (17).

(16)   …: laugh not with him, … least thou *gnash with* thy teeth at the last: …
                                                                    (*EEPF*, 1584)

(17)   …, how hee set his armes one in another, & *nodded with* the head, …
                                                                    (*EEPF*, 1594)

In addition, list (18) specifies an equally large set of examples where, so far, I have been unable to discover any historical or present-day examples using verb-dependent *with*-phrases.

(18)   *bat-eyelid/eyelash/eye, crack-knuckle, flutter-eyelid/eyelash, hunch-shoulder, loll-tongue, roll-eye, shake-head, sway-hip, toss-head, twiddle-thumb/finger, wag-tail, wriggle-toe/ear, wrinkle-nose*, etc.

With some of the collocations in (15), however, the direct object can even today be found to alternate occasionally with the prepositional phrase. Relevant examples include those indicated in bold. Yet here, too, the corresponding proportions of *with*-phrases have generally declined enormously. Consider in Table 2, for instance, the evolution of *stamp (with) one's foot/feet* over the last few centuries, leading, in the databases investigated, from the exclusive occurrence of prepositional phrases to the exclusive use of direct objects. However, as example (19) shows, relevant *with*-phrases associated with the verb *stamp* may still be found today. In particular, this is true of cases like (19) where the verb is separated from the instrumental object by intervening adverbs.

(19)   First it *stamps* vigorously *with* its front feet and …   (*BNC*, file F9F)

To simplify matters, Table 2 disregards the two kinds of context illustrated in (20) and (21).

(20)   … she would stamp her foot upon the ground in a fit of childish impatience, …
                                                                                                          (*NCF*, 1819)

(21)   … and stamping on the ground with her tiny foot, …   (*NCF*, 1830)

Examples like (20), where the body part instrument precedes a prepositional phrase denoting a path or goal, seem to have spearheaded the changeover from the prepositional variant featuring *with* to the direct object. By contrast, (less common) examples like (21), where the body part instrument follows a path or goal expression, may have retained the *with*-variant to this day.

Table 2. Prepositional and direct objects with *stamp (with) one's foot* in a series of historical databases

		I *With*	II Ø	III total	IV percentage of I
1	17th c. (EEPF, EPD)	10	–	10	100%
2	18th c. (ECF, NCF)	16	1	17	94.1%
3	early 19th c. (NCF)	10	–	10	100%
4	late 19th c. to early 20th c. (LNC/B, ETC/B)	14	122	136	10.3%
5	1960s–1993 (wridom1)	–	65	65	0%

2.3   Other instrumental objects

In addition, English boasts a wealth of further instrumental (and related comitative) direct objects alongside traditional *with*-phrases. While some of them may be very old, others have been established in more recent times. Consider, for instance, a representative selection of recent innovations whose translational equivalents in German and other European languages remain restricted to the use of a prepositional phrase or instrumental case form.

(22)   Fishing the floating fly with a very fine cast.   (*OED*, 1922)

(23)   Does the chap shoot a double-barrelled gun?   (*EAF*, 1835)

(24)   a.   …, Bobby Robson decided to play a sweeper … in a 1–2–4–3 formation.
                                                                                        (t90)

       b.   Bobby Robson won't hear of England playing with a sweeper because …
                                                                                        (t90)

As is suggested by (24), the more recent additions to the category of direct object may still use equivalent *with*-phrases on a regular basis.

From the typological perspective, perhaps the most unusual type of instrumental direct objects is found in examples like (25)–(28a), which have been generally neglected so far.

(25)   She wiped the cloth over the dishes. (= She wiped over the dishes with the cloth.)

(26)   She rubbed a towel over his face.

(27)   He … bit his teeth into the succulent meat.
                              <https://archiveofourown.org/works/970505>

(28)   a.   He stroked his hand lightly over her mouth.   (wridom1)
       b.   He stroked lightly over her mouth with his hand.

Example (28a) may be used here to illustrate the two defining properties shared by the examples of this type. In the relevant interpretation, there is a directional element realizing a path or goal function. This means that (28a) does not entail the corresponding sentence without the directional phrase. In addition, the direct object in (28a) corresponds to an instrumental *with*-phrase in a close structural paraphrase such as (28b).

3.   External possessor constructions

The first area where more complex constructions have been phased out at the expense of less complex ones involves examples featuring an inalienable relation between a possessor and a possessum, as in (29a).

(29)   a.   A wasp stung me on the finger.   (wridom1)

As in this example, the ensuing investigation will focus on expressions denoting human beings and (specific parts of) their body. In (29a), the expressions referring to a human possessor and his/her possessum are distributed over two constituents. In (29b), by contrast, they are part of the same constituent.

(29)   b.   A wasp stung my finger.

Accordingly, we can distinguish between external possessor constructions (henceforth, external constructions) as in (29a) and internal ones as in (29b). In the area of external constructions, Old English may be assumed to have possessed a range of patterns abandoned in later stages of English but still preserved in Standard German (for Old English see, e.g., Visser 1963: 627–629, 632, 643–646). I have found it useful, therefore, to compare the situation in English with that in Standard German. In German, there are three major patterns indicated here as A, B, and C and illustrated by the a-examples in (30)–(32).

A.   NP1 – NP2 + NP3

    (30)   a.   *Er massierte mir den Arm.*
           b.   *He massaged me the (my) arm.

B.   NP1 – NP2 + NP3 + PP

    (31)   a.   *Tell legte seinem Sohn einen Apfel auf den Kopf.*
           b.   *Tell put his son an apple on the (his) head.

C.   NP1 – NP2 + PP

    (32)   a.   *Er klopfte mir auf die Schulter.*
           b.   He patted me on the shoulder.

Moving on to the English glosses in the b-examples, we find that (30b) and (31b) are downright unacceptable. In other words, only pattern C seems to be possible in Present-day English. However, in view of fully acceptable examples like (32c), (33), and (34) featuring event nouns (rather than concrete ones), this conclusion needs to be qualified.

(32)   c.   He gave me a pat on the shoulder.

(33)   I turned round and caught him a blow in the ribs.   (wridom1)

(34)   The ball took me an awful whack on the chest.   (*OED*, colloq.)

So what is wrong with (30b) as well as with (31b) is the coincidence of two (adjacent) concrete object NPs. This is why examples such as (32c), (33), and (34) will be disregarded in the rest of this paper. As to pattern C, closer analysis to be presented below has revealed that, unlike German, it is only examples like (32b) containing transitive verbs of surface contact that are productive today. Moreover, in virtually all cases, the sequence S – V – O is entailed by the overall construction S – V – O – PP.

We are now in a position to account for the inclusion of examples like (25)–(28a) and (29a) or (32b) in Section 2.3 and Section 3, respectively. It is true that all of these examples instantiate the frame S – V – O – PP and that the prepositional phrase constitutes an indispensable feature of the construction. However, while it is only natural to present examples like (25)–(28a) along with others involving instrumental objects, it would also seem to be appropriate to treat examples of pattern C as in (29a) or (32b) along with more complex types of external possessor constructions.

More importantly, pattern C (unlike cases like (25)–(28a) in the preceding section) involves an additional and inalienable relation between an external possessor and a possessum. Arguably, this feature may be assumed to make pattern C more complex than examples (25)–(28a). As suggested by Visser (1963: 643), the NP in the possessum phrase "is virtually an object: in 'he patted the boy on the shoulder' both the boy and the shoulder are patted". Moreover, the classification adopted is fully in line with the overall results of this study: while instrumental objects as in (25)–(28a) appear to be increasingly productive, pattern C has been declining for centuries in its functional range.

In the following, I want to show that patterns A and B, which arguably may be regarded as more complex than pattern C, also occurred, if only marginally, in Early and Late Modern English. In addition, it will be demonstrated that the decline of external constructions has also affected pattern C.[3] Concerning pattern A, I have

---

3.   Restrictions of space prevent me from detailing a range of issues found with external constructions and their internal counterparts. As regards the possessum phrase in external constructions, they include contrasts between the definite article and the possessive pronoun, the singular and the plural as well as between prepositions (e.g. *in* vs *into*). Concerning internal constructions, no distinction is made in the relevant counts between the three means of indicating the possessor, possessive pronouns, genitives, and postposed *of*-phrases. In addition, this study neglects the use of manner adverbs modifying the verb phrase.

so far only been able to detect three kinds of relevant examples involving specific verbs and nouns denoting body parts or personal attributes. The first two are seen in a quotation provided by the *OED*:

(35)   Those holie martyrs, whom the Emperor Maximus had put out the right eie, and hockt their left legs.   (*OED*, 1563–87)

Notice that in both examples of (35) possessor and possessum have been separated by the relativization of the possessor phrase. As Table 3 shows, in all other uses of *put out* associated with the noun *eye* in the *OED* quotations between 1297 and 1659, the internal construction has been selected instead. Thus it is likely that it is the extraction context in (35) which is responsible for the occasional use of such combinations well into the 16th century. More specifically, I would submit that the longer survival of pattern A in such environments is primarily due to the backgrounding of the relationship holding between the two object NPs. In addition, the survival of cases like (35) would also be in line with Keenan and Comrie's Accessibility Theory (Keenan & Comrie 1977), which stipulates that objects rank higher than genitives (Hendrik De Smet p.c.). Similar tendencies avoiding the adjacency of two objects in other than such external constructions will be pointed out in Section 4.

Table 3. External and internal constructions with the verb *put out* and the noun *eye* in the *OED* quotations

		I external	II internal
1	1297–1598	1 (1563–87)	11
2	1606–1659	–	13

The third manifestation of type A does not refer to people and their body parts, and, therefore, the presence of inalienable possession might be in doubt. Here the verb *cross* in (36) is associated with the noun *way*, which I interpret as belonging to the personal sphere of the person in question.

(36)   This man..if but an hare crosse him the way, he returnes.   (*OED*, 1608)

The few examples I have found so far date from the late 16th and early 17th centuries. They contrast with examples like (37), which are much more common in the Early Modern period and whose status as the corresponding internal construction might likewise be questioned.

(37)   What if a Hare crosse your way?   (*EPD*, 1613)

This takes us to pattern B, which was only found with two verbs of whispering associated with the possessum phrases *in the/his ear*, as in examples (38a) and (39a).

(38) a. …, hee rounds Sebastiano this answer in his eare, …   (*EEPF*, 1635)

(39) a. … when a Messenger … whisper'd them something in the Eare, …

(*EEPF*, 1655)

The corresponding examples of the simpler pattern C not containing a patient object are exemplified in (38b) and (39b).

(38) b. … she rownde her father in the eare, and sayde: …   (*EEPF*, 1535)

(39) b. …; when one who sate by him whisper'd him in the Ear, and …

(*EEPF*, 1555)

Consider now in Tables 4 and 5 the evolution of the two related constructions involving the near-synonymous verbs *round* and *whisper*, respectively. Both tables supply two kinds of information. Columns I and II compare the aggregate figures of the two external constructions in (38a–b) and (39a–b) with their corresponding internal ones. In addition, the bracketed figures in column I contrast the a-examples involving an object additional to the possessor NP (pattern B) and the b-examples containing only the possessor NP (pattern C). In column II, a corresponding distinction is made between the internal constructions with or without an additional object phrase. The internal constructions for the two verbs are illustrated in (40) and (41a–b). However, in the case of *round*, the internal construction lacking a direct object is not met with in the corpora consulted.

(40) This he was continually rounding in my ears.   (*ECF*, 1739)

(41) a. …, when Queen Elizabeth whispered these Words in his Ear.

(*EPD*, 1691)

b. Another whispered in her Ear, …   (*ECF*, 1742)

Table 4. External and internal constructions involving the verb *round* and the noun *ear* in a series of historical databases

		I external *x in* *the ear* (B vs. C)	II internal *in(to)* *x's ear*	III total	IV percentage of I (B vs. C)
1	16th–17th c. (EEPF, EPD)	22 (2/20)	–	22 (2/20)	100%
2	18th c. (ECF, EPD, NCF)	–	1 (plural) (1/0)	1 (1/0)	
3	19th c. (NCF, EPD)	–	1 (1/0)	1 (1/0)	
4	1960s–1993 (BNC: overall)	–	–	–	–

Table 5. External and internal constructions involving the verb *whisper* and the noun *ear* in a series of historical databases[*]

	I external *x in* *the ear* (B vs. C)	II internal *in(to) x's ear*	III total	IV ratio of I (B vs. C)
1 16th–17th c. (EEPF)	56 (4/52)	64 (30/34)	120 (34/86)	46.7% (13.3%/60.5%)
2 18th c. (ECF, EPD, NCF)	18 (0/18)	81 (36/45)	99 (36/63)	18.2% (0%/28.6%)
3 19th c. (NCF, EPD)	–	139 (45/94)	139 (45/94)	0%
4 1960s–1993 (BNC: overall)	–	143	143	0%

[*] The analysis includes a small number of relevant passive examples.

On the basis of the evidence in Tables 4 and 5, we can make a number of observations.

– Compared with pattern C, which was still common with at least *whisper* up to the 18th century, pattern B was even in the Early Modern period a rare phenomenon, and with both verbs it seems to have been phased out before the end of the 17th century.

– In the case of *round*, which lacked the support of corresponding internal constructions, the simpler pattern C did not survive beyond the 17th century either.

– With *whisper*, however, pattern C is found as late as the 18th century. Moreover, unlike *round*, internal constructions with *whisper* corresponding to both patterns B and C (i.e. with or without an additional patient object) survive to this day.

Apart from *whisper* and *round*, three less common verbs of whispering – *whister, whistle*, and *buzz* – representing pattern C and associated with the possessum phrase *in the ear*, have also been attested for Early Modern English. Consider, for instance:

(42)   …; but for more assured proufe hee whistered the kinge in the eare desiring him that …   (*EEPF*, 1576)

All of these uses have become obsolete today. There are a number of further manifestations of pattern C that have by now dropped out of use. To begin with, there were examples like (43), which contain the verb *run* and prepositional phrases referring to various parts of the body or the noun *body* itself. Such examples are found as late as the 19th century.

(43)   … he twice ran him thorow the body with his sword.   (*EEPF*, 1635)

There was no internal construction directly corresponding to cases like (43) or those omitting the instrumental *with*-phrase.

The majority of obsolete cases representing pattern C are characterized by the presence of *in the face* as the possessum constituent. The external constructions in question include five basically intransitive verbs (*gaze, glare, grin, smile, sneer),* as illustrated in examples (44) and (45).

(44)   … she gazed him full in the face, …   (*NCF*, 1861)

(45)   …; The Injury he had done his Wife and Family now glared him full in the
       Face, …   (*ECF*, 1741)

An overview of the replacement of the external construction by the internal one is
given in Table 6. It is apparent that after the Early Modern period there is a striking
decline of pattern C.[4]

Table 6. External and internal constructions involving the verbs *gaze, glare, grin, smile,*
and *sneer* and the noun *face* in a series of historical databases (*EEPF, EPD, ECF, NCF,*
wridom1)

		I external *x in* the face	II internal *in(to)* *x's face*	III total	IV ratio of I
1	16th–17th c.	10	19	29	34.5%
2	18th c.	1	45	46	2.2%
3	19th c.	2	62	64	3.1%
4	1960s–1993	–	72	72	0%

Table 6 omits the verb *laugh*, which American informants assure me could still be
used today in external constructions. The historical databases used in this article
provide only two further examples of this pattern, with both of them displaying
inanimate subjects and non-literal (or idiomatic) readings as in (46).

(46)   …, his barns are full, and generous plenty laughs him in the face. (*EAF*, 1849)

Moreover, out of a total of five relevant attestations for *glare*, the two most recent
ones dating from the 18th and 19th centuries also involve non-literal uses associ-
ated with inanimate subjects, as in (45).

As mentioned before, it is only transitive verbs of surface contact like *sting*
and *pat* as in (29a) and (32b), respectively, that can be said to be productive today
in external constructions of pattern C. Thus, they contrast with three basically
intransitive verbs, *look, stare,* and *laugh*, which also occur in such examples. While
relevant examples of the transitive type may be regularly reduced to the sequence
S – V – O, which they entail, the intransitive type involving *look, stare,* and *laugh*
cannot omit the possessum phrase. When used in pattern C, the three verbs are

---

4. The Early Modern English database attests two related transitive verbs, *behold* and *view*, which
also occurred, though very rarely, in external constructions of pattern C featuring the possessum
phrase *in the face*. For instance:

(i)   He that had that time beholden the Baron in the face, would have thought …
                                                                               (*EEPF*, 1567)

usually associated with the possessum phrase *in the face*, which, unlike the situation in productive transitive verbs, cannot be modified.[5] Owing to the infrequency of relevant examples using *laugh*, the following discussion will focus on *look* and *stare*, which are still regularly found in external constructions. However, closer analysis of their behavior has revealed a centuries-old decrease of relevant basic uses. In the case of *look*, we can observe a massive replacement of the external construction in examples like (47a) by the corresponding internal one as in (47b).

(47) a.   She kept looking him in the face.
      b.   She kept looking in(to) his face.

Table 7 shows the dramatic decline of the external construction between the Early Modern period and the present day.

Table 7. External and internal constructions involving the verb *look* and the noun *face* in the 16th to17th centuries and the late 20th century

		I	II	III	IV
		x in the face	*in(to) x's face*	total	percentage of I
1	16th–17th c. (EEPF)	56	21	77	72%
2	1960s–1993 (wridom1)	63	147[*]	210	30%

* With two exceptions, the earlier linking preposition *in* has here been replaced by *into*.

Incidentally, the modest role played by examples like (48) featuring non-literal readings associated with inanimate object NPs does not seriously impact on the results of this analysis.

(48)   I shall … sweat it out and look the facts in the face, …   (wridom1)

The evolution of the verb *stare* is decidedly more complicated. On the one hand, straightforward literal examples as in (49) and Table 8, where both the subject and the possessor object refer to human beings, have experienced an equally dramatic decline at the expense of the internal possessor construction.

(49)   Mrs Bennett … stared her straight in the face.   (wridom1)

Table 8. Comparing straightforward literal external and internal constructions involving the verb *stare* and the noun *face* in the 16th to17th centuries and the late 20th century

		I	II	III	IV
		x in the face	*in(to) x's face*	total	percentage of I
1	16th–17th c. (EEPF)	14	10	24	58.3%
2	1960s–1993 (wridom1)	5	42	47	10.6%

5.   With *look*, there is the alternative *in the eye*, which became established in the 19th century.

On the other hand, the external construction has increasingly adopted the non-literal orientations shown in (50)–(52), which tend to be associated with inanimate subjects and possessor NPs. Remember that the trend towards idiomatic readings was also noted above in connection with the verbs *laugh* and *glare*.

(50) Ruin stared us in the Face.   (*ECF*, 1765)

(51) It has been staring us in the face for months and we never even twigged.

(wridom1)

(52) He had stared death in the face and …   (wridom1)

Table 9 specifies the relevant proportions of such non-literal uses. It is apparent that over the last few centuries the incidence of these uses has increased from 12.5% to over 80%. In the corresponding internal construction, by contrast, similar idiomatic uses have always been extremely rare.[6]

Table 9. External and internal constructions involving the verb *stare* and the noun *face* in a series of historical databases[*]

		I 16th–17th c. (EEPF)	II 18th–early 19th c. (ECF, NCF1)	III *1830–*1869 (LNC/B)	IV 1960s–1993 (BNC: wridom1)
1	*x in the face*	12.5% (*n* = 16)	51.1% (*n* = 43)	81.6% (*n* = 38)	82.8% (*n* = 29)
2	*in(to) x's face*	0% (*n* = 10)	7.4% (*n* = 27)	7.1% (*n* = 14)	2.3% (*n* = 43)

* The percentages indicate the ratios of (stereotyped) non-literal uses as in examples (42)–(44).

Does the decline of (literal) external constructions extend to basically transitive verbs? So far, the issue has been explored by analyzing only two collocations featuring *by*-phrases as their possessum constituents, *take x by the arm* and *shake x by the hand*, as illustrated in (53)–(54).

(53) She felt Matey take her by the arm and …   (wridom1)

(54) Now he … shook him formally by the hand.   (wridom1)

Consider now the analyses summarized in Tables 10 and 11.

---

6. Unlike German, where non-literal readings tend to be confined to external constructions, English generally uses internal constructions to convey both literal and non-literal readings (König & Gast 2007: 199–200). As we have seen just now, there is a residue of cases involving *stare*, *look*, and possibly *laugh*, which English uses more or less widely in the external construction to convey non-literal readings as well. In addition, centuries-old non-literal fixed expressions like *look a gift horse in the mouth* (which is entirely unacceptable in a literal reading) and *hit the nail on the head* suggest that the tendency to associate external constructions with non-literal readings must have been more widespread in former periods.

Table 10. Comparing the types *he took x by the arm* and *he took x's arm* in the 16th to 17th centuries and the late 20th century

		I *x by the arm*	II *x's arm*	III total	IV percentage of I
1	16th–17th c. (EEPF)	31	2	33	93.9%
2	1960s–1993 (wridom1)	59	241	300	19.7%

Table 11. Comparing the types *he shook x by the hand* and *he shook x's hand* in the 16th to 17th centuries and the late 20th century

		I *x by the hand*	II *x's hand*	III total	IV percentage of I
1	16th –17th c. (EEPF)	10	–	10	100%
2	1960s–1993 (wridom1)	27	154	181	14.9%

Clearly, the evidence leaves no doubt that here, too, the last few centuries have witnessed a dramatic decline of external possessor constructions at the expense of internal ones.

4.   Double object constructions

Since Old English times, English has also experienced striking reductions in a second area of complex constructions, that of double objects. The verb classes which – with few exceptions – have given up the double object construction over the last few centuries include the following (Rohdenburg 1995a: 109–112):

(55)   a.   verbs of separation: *banish, bar, discard, discharge, dismiss, eject, exclude, exile, expel, expulse*, etc.
  b.   verbs of dispossession: *(be)reave, defraud, deprive, dispossess, rob*, etc.
  c.   directive verbs: *advise, ask, beseech, bid, command, counsel, enjoin, persuade, pray, request, warn*, etc.

Such uses, no longer acceptable today, are illustrated in (56)–(58).

(56)   This said, the king in a fury returns to the Palace, and banished Dampion the Court, …   (*EEPF*, 1665)

(57)   Can Children esteeme this for tender love, which deprives them their sight whom they only love?   (*EEPF*, 1640)

(58)   I was going on to more Particulars, when my Master commanded me Silence.   (*ECF*, 1726)

Discounting the alternative use of clausal constructions, these changes have generally resulted in two kinds of increased structural explicitness:

i.   Double objects have largely given way to less opaque expressions consisting of (direct) object + prepositional phrase.
ii. By narrowing down the functional range of the double object construction, English has – *pace* Hawkins (1986) – evolved a much closer form-function fit in the remaining double objects.

Thus the second objects associated with the verbs in (55a) and (55b) have been regularly replaced by more explicit prepositional phrases including the items *from, of*, and *out of*. Compare, for instance:

(59)   … that he should be expelled from the university …   (*ECF*, 1751)

By contrast, in the case of directive verbs in (55c), the less common double objects have only in some cases been continued by the combination of object + *to/into/ toward(s)*-phrases, as in (60).

(60)   … she would not enjoin her messenger to secrecy …   (*NCF*, 1811)

The evolution of the directive verb *command* summarized in Table 12 would seem to be typical of the dramatic decline and eventual demise of the double object experienced by the verbs in (55c).

Table 12. The decline of double objects associated with the directive verb *command* between the 17th and the late 20th century[*]/[**]

		I number of examples	II size of the database	IV frequency per million words
1	17th c. (EEPF)	47 (14/33)	6,458,143	7.277
2	18th c. (ECF)	8 (4/4)	9,702,696	0.414
3	late 18th to mid 19th c. (NCF)	1 (1/0)	37,415,696	0.026
4	late 19th to early 20th c. (LNC/B, ETC/B)	4 (1/3)	25,619,210	0.156
5	1960s–1993 (BNC: wridom1)	–	18,863,529	–

[*] The analysis is confined to personal pronouns (excluding *it*) functioning as experiencers immediately following the verb.
[**] The bracketed figures distinguish between examples where the object following the experiencer has been extracted and all others.

Interestingly, several of the verbs in (55a) and (55c) have been found to display a definite tendency to separate, and thus to mitigate, the coincidence of two nominal arguments during the final stages of their existence. This can be achieved in two ways. First, the verbs of separation in (55a) usually tend to preserve the double object construction much longer in the passive, where the two NPs have to be separated as in (61) (Rohdenburg 1995a: 109–112). With passive *dismiss*, and unlike with the active, such examples are still found in British English today (Rohdenburg 2009a: 202–203).

(61)   The commander was formally dismissed the service, …   (t94)

Second, in the active, the separation of the two objects may be brought about by various types of extraction, as in (62).

(62)   … a resolute denial of what the great Mithridates has commanded you, …
                                                                              (*EEPF*, 1669)

The bracketed figures in Table 12 show that such devices were regularly resorted to for this purpose. The phenomenon is reminiscent of example (35), where it is the relativization of the possessor NP that avoids the adjacency of possessor and possessum.

In general, the English double object construction has been confined to the verbs of transfer and communication as well as a few allied verbs, thus producing a much closer form – function fit in the remaining examples. But even in this class of verbs, the trend towards narrowing the semantic spectrum of the double object construction has continued unabated over the last few centuries. Thus several studies have shown that double objects have become rare, obsolescent, or obsolete with a number of verbs like those in (63a–b) (see, e.g., Rohdenburg 2009a, 2009b; Colleman 2011; Colleman & De Clerck 2011).

(63)   a.   *answer, deliver, restore, return*
         b.   *furnish, present, provide, supply*

In illustration, consider in Tables 13 and 14 the decline of the double object construction with the verbs *deliver* and *provide*.

Table 13. Double objects involving the verb *deliver* and patient NPs containing the noun *letter* in a series of historical databases[*]

		I double object	II presence of only the patient object	III total	IV percentage of the double object
1	17th c. (EEPF)	70	107	177	39.5%
2	18th c. (ECF)	28	90	118	23.7%
3	*1728–*1799 (NCF1)	5	59	64	7.8
4	*1800–*1869 (NCF2, LNC/B)	2	64	66	3.0
5	1960s–1993 (BNC overall)	–	20	20	0%

[*] The analysis is confined to personal pronouns (excluding *it*) immediately following the verb.

Present-day British English seems to have phased out by now examples like (64), where agentive subjects co-occur with concrete objects. However, the double object survives in abstract uses of the verb involving, in particular, examples such as (65), where the subject refers to an inanimate entity. Similar phenomena of detransitivization in the sense of Hopper & Thompson (1980), which help an otherwise

precarious double object construction to survive, are also found with other verbs including *present*.

(64)   Mrs. Clinton delivered Madame Duval a letter from Mr Villars, which …
                                                                                    (*ECF*, 1778)

(65)   His service, which has delivered him an astounding 957 aces this year, looked
         impregnable.   (t92)

Table 14. Double objects and sequences of the type object + *with*-phrase associated with
the verb *provide* in a series of historical databases[*]

		I double object	II object + *with* +NP	III total	IV percentage of the double object
1	[*]1660–[*]1699 (ECF)	34 (33/1)	13 (13/0)	47 (46/1)	72.3%
2	[*]1740–[*]1799 (NCF)	17 (13/4)	26 (21/5)	43 (34/9)	39.5% (38.2%/44.4%)
3	[*]1800–[*]1829 (NCF)	7 (1/6)	48 (41/7)	55 (42/13)	12.7% (2.4%/46.2%)
4	[*]1830–[*]1869 (NCF,/LNC/B)	2 (0/2)	100 (94/6)	102 (94/8)	1.9% (0%/25%)
5	1960s–1993 (wridom1)	7 (4/3)	132 (127/5)	139 (131/8)	5.0% (3.1/37.5)

[*] The analysis is confined to personal pronouns (excluding *it and 'em*) immediately following the verb. The bracketed figures distinguish between the canonical ordering and those in which the direct object or the NP of the prepositional object have been extracted by relativization or some other means.

In Table 14, the division of the overall historical totals of double objects and NP1 + *with* + NP2 into canonical orderings and extracted second NPs corroborates the findings in Rohdenburg (2009a, 2009b) concerning *present* and *furnish*, respectively: the second NP of the double object construction is proportionally more likely to be extracted than that of the sequence NP1 + *with* + NP2. In other words, the double object has been preserved much better in those cases where the two objects have been separated from each other. However, we cannot exclude the possibility that the phenomenon serves more than one purpose, not just the enhancement of the recessive double object but also the avoidance of preposition stranding as referred to in Yañez-Bouza (2014) and Rohdenburg (2017). Moreover, as pointed out by Hendrik De Smet, it is in in line with Keenan and Comrie's (1977) Accessibility Hierarchy, which predicts extraction from a prepositional phrase to be harder than extraction of an object.

5.   Clausal complements after the sequence V+O

Clausal complements after transitive verbs (i.e. those associated with a direct object) constitute a third area of complex constructions which has experienced a sweeping restructuring over the last few centuries.

5.1   *That*-clauses

In particular, it is the frame direct object + *that*-clause whose range of application has contracted enormously. In the following, we will take a brief look at three verb classes that have either completely or in part given up the possibility of being complemented by a *that*-clause in this environment. It may be assumed that the verbs of negative causation in (66) were the first to abandon the *that*-clause, as in (67), and to adopt instead the gerund, which is typically introduced by *from* as in (68).

(66)   verbs of negative causation: *ban, bar, discourage, disqualify, forbid, hinder, inhibit, prevent, prohibit, refrain, restrain, deter, withhold*, etc.

(67)   They scarse refrained the people, that they did not sacrifice vnto them.

<div align="right">(OED, 1535)</div>

(68)   …, yet could he not refrain his eyes from beholding her, …   (*EEPF*, 1576)

The assumption is based on an informal comparison of the directive and expressive verbs (Rohdenburg 1995b and 2014, respectively) with the verbs *forbid, prohibit, prevent*, and *hinder*, as found in the *EEPF*, the *EPD*, and the *OED* quotations database (concerning the latter, see also Iyeri 2010). The gerund, which is introduced by the semantically motivated item *from* as an obligatory or potential feature, has been around for centuries (see, e.g., Iyeri 2010: passim). Crucially, however, and unlike the gerund, only a small minority of the *that*-clauses associated with these verbs in the Early Modern English period are preceded by a direct object. This suggests that the *that*-clause survived much longer in corresponding intransitive uses (i.e. those not involving a direct object).

   Equally dramatic, if somewhat later, changes have occurred in the area of the directive verbs in (69), where the *that*-clause as in (70) has been superseded by the *to*-infinitive as in (71) (Rohdenburg 1995b).

(69)   directives: *admonish*, **advise, beg,** *beseech, bid, charge*, **command, counsel, direct, enjoin, entreat, instruct**, *persuade*, **pray**, *order. recommend. request*, etc.

(70)   … the Priest of the God commaunded them that they should not presume to read it, before …   (*EEPF*, 1588)

(71)   At which words, ye knight commanded them to cease their crueltie, …

<div align="right">(EEPF, 1598)</div>

This means that transitively used directive verbs have over time given up two older types of complex object construction, *that*-clauses and double objects (illustrated in Section 4). The evolution of the *that*-clause after the transitive use of directive verbs between the16th and 18th centuries is summarized in Table 15, which is based on Table 3 in Rohdenburg (1995b: 372–373).

Table 15. Finite complement clauses and infinitives after the combination directive verb + object in seventeen 16th to18th century narrative authors[*]

	I *that*-clause	II infinitive	III total	IV percentage of the *that*-clause
1 16th c. (Roper c1540 – Nashe 1594)	21	27	48	43.8%
2 17th c. (Wilkins 1608 – Bunyan 1684)	27	138	165	16.4%
3 18th c. (Defoe 1719 – Bligh 1790)	25	c. 250	c. 275	9.1%

* While the vast majority of the occurrences retrieved involve direct objects, the analysis includes one or two verbs that use the occasional prepositional object along with the direct one (e.g. *entreat (of)*).

It is apparent that there is a steady decrease of the proportion of *that*-clauses at the expense of infinitival complements, and by the beginning of the 19th century, the *that*-clause seems to have reached its final phase. Significantly, the outgoing type of complement is not found in the – early 19th century – novels by Jane Austen analyzed in Rohdenburg (1995b). With very few exceptions, the *that*-clause after directive verbs has now been replaced completely by the *to*-infinitive.[7] By contrast, any existing intransitive uses of these verbs (shown in bold in (69)) still permit the use of the *that*-clause.

　　Similar, though less far reaching, reductions of the *that*-clause have affected the class of expressive verbs (Rohdenburg 2014). These verbs have evolved a rival complement in a prepositional gerund, which has largely replaced the *that*-clause, as in (72a).

(72)   a.   They praised her for being so unselfish.
　　　　b.   They praised her for her unselfishness.

As is suggested by examples like (72a–b), the emergence of the gerunds follows in the wake of corresponding prepositional phrases: The gerund typically selects the same preposition that is found with nominal complements. The list in (73) includes some of the verbs where the prepositional gerund has generally superseded the *that*-clause.

(73)   expressives: *accuse, curse, excuse, praise, rebuke, thank, upbraid*, etc.

7.   Admittedly, while the subject of the *that*-clause and the object of the superordinate verb may – in extremely rare cases – be found to be referentially distinct, the implied subject of the infinitive is necessarily co-referential with the object of the matrix clause.

However, there is a small residue of verbs listed in (74) which can still be found with the traditional *that*-clause (mostly alongside the more recent gerund).

(74) a. *pride, flatter*
      b. *blame, congratulate, reproach*, etc.

In this case, *flatter* and *pride* invariably select reflexive pronouns, and even the remaining verbs in (74b) are typically associated with reflexive objects and other manifestations of Low Transitivity in the sense of Hopper & Thompson (1980). Examples illustrating such functional niches permitting the survival of the *that*-clause in Present-day English are presented in (75) and (76).

(75) The president can congratulate himself that so far he has not, as he put it, made the wrong mistakes.   (t90)

(76) We shall not be blaming ourselves that what we said did not square with what the voters did.   (t92)

This brings us to the class of representative verbs in (77) (Rohdenburg 2014).

(77) representatives: **acquaint**, *advise*, **answer**, *assure*, **certify**, **confirm**, *instruct, let x know, notify, remind, satisfy, teach, tell, warn*

The majority of these verbs have retained the use of the sequence direct object + *that*-clause. It is only the four verbs in bold that have abandoned the frame altogether. In illustration, consider such obsolete uses with *answer, certify*, and *confirm*:

(78) I answered him that I had found a friend, whose principles …   (*NCF*, 1794)

(79) Plangus certified them that all things were prepared for their accommodation, …
                                                                          (*ECF*, 1651)

(80) I have confirm'd him that you have not such a thought, …   (*EPD*, 1662)

Crucially, however, the verbs illustrated in (78)–(80) continue to select the *that*-clause in – originally available – intransitive uses.

    In this connection, the case of *satisfy*, which cannot be used intransitively preceding the *that*-clause, is particularly revealing. Here, the sequence object + *that*-clause tends to display a special affinity for two manifestations of Low Transitivity, reflexive objects and/or inanimate subjects. It is shown in Rohdenburg (2014: 168–169) that – for two centuries – the relevant use of *satisfy* has been associated with the two environments in at least 90% of all cases.

5.2   Infinitival complements

As mentioned above, the transitively used directive verbs have usually abandoned the *that*-clause in favor of infinitival complements. By contrast, most other verb classes have increasingly eliminated the object + infinitive combination. Discounting the exceptional survival of this frame with *forbid*, the verb classes that have phased out this option include the following (Rohdenburg 1992):

(81) a. verbs of negative causation as in (82)
      b. expressive verbs as in (83)
      c. representatives as in (84)

(82) …; that she cannot prevent any Gentleman to write and send her a letter, …
                                                        (*EEPF*, 1635)

(83) The Bishops charged the Protestants to have been the propounders of the questions.   (*OED*, 1709)

(84) …, he arising from his seate, and taking Brusanus in his armes assured him to be as heartelye welcome to the King Leonarchus, as euer he had bene before to the Marchaunt Corynus: …   (*EEPF*, 1592)

Furthermore, in the class of commissive verbs including *threaten, offer, proffer*, and *promise*, only the last one is still found in some varieties to use the subject-controlled infinitive in the frame S – V – O + infinitive as a recessive alternative to the *that*-clause (Rohdenburg 1992: 201–202). With the rest, infinitives controlled by both the subject and the object as in (85) and (86), respectively, seem to have fallen out of use.

(85) I'd threaten him to put myself into a Nunnery.   (*EPD*, 1739)

(86) And did your father proffer you to be his wife.   (*EPD*, 1653)

In other words, discounting the delayed case of *promise* and so-called subject-to-object raising structures, the combination of direct object + infinitive has been virtually confined to the class of directive verbs. By contrast, the (subject-controlled) infinitival construction has survived in the simpler intransitive counterparts of examples like (85).

5.3   Dependent interrogative clauses

English has shown for centuries a tendency to mark dependent interrogatives by means of (verbal and) prepositional links, as in (87), thus aligning the use of prepositional phrases containing NPs with that of subordinate interrogative clauses (see, e.g., Rohdenburg 2003).

(87)   a.   We consulted (on) how I should proceed.
       b.   He consulted us (on) how he should proceed.

Crucially, these changes are typically speeded up after transitive verbs, as shown in Table 16 (based on Table 10 in Rohdenburg 2003: 218). No doubt, the inclusion of a (syntactically and semantically motivated) preposition in examples like (87a–b) has increased their grammatical explicitness.

Table 16. Prepositional and zero links with interrogative clauses dependent on the verb *consult* in databases spanning the 18th and 19th centuries (*ECF, EAF, NCF*)

		I added prepositions	II Ø	III total	IV ratio of added prepositions
1	intransitive uses	29	133	162	17.9%
2	transitive uses	14	11	25	56%

6.   Conclusion

We have seen, then, that in simple structures English displays an ongoing trend towards transitivization, which has led to the replacement of a vast range of prepositional objects and adjuncts by direct objects. In line with Hawkins' drift theory, this has undoubtedly made the category of direct objects more diverse or general in functional terms.

On the other hand, a whole range of more complex object constructions have been noticeably restricted or phased out altogether. In the process, the form–function fit, the mapping between form and meaning, has in general been enhanced considerably. This is true of most of the complex structures discussed in Sections 3–5.

In Section 3, we find that the domain of external possessor constructions has been severely restricted at the expense of corresponding internal ones. While the latter are structurally simpler than their external counterparts, the overall reductive changes may be argued to have produced an enhanced form–function fit. At any rate, the only surviving frame in the domain of external possessor constructions, S – V – O – PP (pattern C), has become extremely specialized.

Similarly, in Section 4, dealing with the area of double objects, the range of compatible verbs has been strikingly reduced. As a result, the construction has been phased out completely with several verb classes, and even the remaining stronghold, the verbs of transfer and communication, has been drastically restricted. Crucially, by narrowing the semantic spectrum, these developments have produced a high degree of specialization in this area.

Finally, as part of the so-called Great Complement Shift (e.g. Rohdenburg 2006; Vosberg 2006; Rudanko 2015), complement clauses involving the frames S – V – O + *that*-clause and S – V – O + *to*-infinitive, which used to cover a broad semantic spectrum, have experienced a series of major changes. In the process, several verb classes associated with these frames have either completely or in part dropped out of use, and their functions have to some extent been taken over by novel prepositional gerunds. As a result, the three major syntactic frames, S – V – O + *that*-clause, S – V – O + *to*-infinitive, and S – V – O + prepositional gerund, have evolved a degree of specialization unknown to languages like German.

Comparing now the trend towards functional diversity displayed by the direct object in simple structures and the semantic specialization characteristic of complex object constructions, we might be tempted to regard these antagonistic tendencies as a kind of trade-off between the amount of processing required per syntactic frame and the degree of functional diversity permitted. This interpretation is supported by the fact that in a number of areas English – unlike languages such as German – distinguishes sharply between clausal complements with or without a preceding direct object. This holds for *that*-clauses and to a lesser extent also for infinitival complements. Unlike combinations of direct object + *that*-clause that have been phased out, originally available *that*-clauses without a preceding object have generally survived intact. So have infinitives after intransitively used commissive verbs, whose transitive counterparts have fallen out of use. In Section 5.3, a similar contrast has been noted in the domain of dependent interrogative clauses, which in recent times have begun to be introduced by explicit prepositional linkers. However, here, too, we can observe an evolutionary contrast between transitive and intransitive matrix verbs. Significantly, and in line with the Complexity Principle (see, e.g., Rohdenburg 2003, 2009b), it is the more complex examples containing transitive verbs that are leading the trend towards greater structural explicitness.

In addition, we have observed that the decline of complex constructions may be accompanied (and possibly foreshadowed) by three more or less general trends:

1. the tendency of several verbs to separate two objects by means of passivization and object extraction,
2. the increasing restriction of the sequence object + *that*-clause to Low Transitivity environments (reflexive objects, inanimate subjects and the passive), and
3. the decline of literally interpreted constructions in favor of non-literal (or idiomatic) ones in the case of external possessor constructions.

References

Bibovič, Ljiljana. 1976. On the notion of body part instrument. *Folia Linguistica* 9: 311–324. doi:10.1515/flin.1976.9.1-4.311

Colleman, Timothy. 2011. Ditransitive verbs and the ditransitive construction: A diachronic perspective. *Zeitschrift für Anglistik und Amerikanistik* 59: 387–410.   doi:10.1515/zaa-2011-0408

Colleman, Timothy & De Clerck, Bernard. 2011. Constructional semantics on the move: On semantic specialization in the English double object construction. *Cognitive Linguistics* 22: 183–209.   doi:10.1515/cogl.2011.008

Hawkins, John A. 1986. *A Comparative Typology of English and German: Unifying the Contrasts*. London: Croom Helm.

Hopper, Paul J. & Thompson, Sandra A. 1980. Transitivity in grammar and discourse. *Language* 56: 251–299.   doi:10.1353/lan.1980.0017

Iyeiri, Yoko. 2010. *Verbs of Implicit Negation and their Complements in the History of English*. Amsterdam: John Benjamins.   doi:10.1075/z.155

Jespersen, Otto. 1927. *A Modern English Grammar on Historical Principles, Part III: Syntax*, Vol. 2. London: George Allen & Unwin and Copenhagen: Ejnar Munksgaard. Reprinted 1961 and 1965 by Dickens & Co., Northampton.

Keenan, Edward L. & Comrie, Bernard. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63–99.

Kirchner, Gustav. 1955. Direct transitivation. *English Studies* 36:15–23. doi:10.1080/00138385508596940

Kirchner, Gustav. 1957. Recent American influence on standard English: The syntactical sphere. *Zeitschrift für Anglistik und Amerikanistik* 5: 29–42.

Kirchner, Gustav. 1959. Zur transitiven und intransitiven Verwendung des englischen Verbums. *Zeitschrift für Anglistik und Amerikanistik* 7: 342–399.

König, Ekkehard & Gast, Volker. 2007. *Understanding English-German Contrasts*. Berlin: Schmidt.

Legenhausen, Lienhard. 1988. Prototypical lexical causatives and transitional relations to non-causatives. In *Essays on the English Language and Applied Linguistics on the Occasion of Gerhard Nickel's 60th Birthday*, Dietrich Nehls (ed.), 131–146. Heidelberg: Groos.

Plank, Frans. 1983. Transparent versus functional encoding of grammatical relations: A parameter for syntactic change and typology. *Linguistische Berichte* 86: 1–13.

Plank, Frans. 1984. Verbs and objects in semantic agreement: Minor differences between English and German that might suggest a major one. *Journal of Semantics* 3: 305–360. doi:10.1093/jos/3.4.305

Rohdenburg, Günter. 1974. *Sekundäre Subjektivierungen im Englischen und Deutschen: Vergleichende Untersuchungen zur Verb- und Adjektivsyntax*. Bielefeld: Cornelsen, Velhagen & Klasing.

Rohdenburg, Günter. 1992. Bemerkungen zu infiniten Konstruktionen im Englischen und Deutschen. In *New departures in contrastive linguistics/Neue Ansätze in der Kontrastiven Linguistik*, Vol. I, Christian Mair & Manfred Markus (eds), 187–207. Innsbruck: Universität Innsbruck.

Rohdenburg, Günter. 1995a. Betrachtungen zum Auf- und Abstieg einiger präpositionaler Konstruktionen im Englischen. *NOWELE* 26: 67–124.   doi:10.1075/nowele.26.05roh

Rohdenburg, Günter. 1995b. On the replacement of finite complement clauses by infinitives in English. *English Studies* 76: 367–388.   doi:10.1080/00138389508598980

Rohdenburg, Günter. 2003. Cognitive complexity and *horror aequi* as factors determining the use of interrogative clause linkers in English. In *Determinants of Grammatical Variation in English*, Günter Rohdenburg & Britta Mondorf (eds), 205–249. Berlin: Mouton de Gruyter. doi:10.1515/9783110900019.205

Rohdenburg, Günter. 2006. The role of functional constraints in the evolution of the English complementation system. In *Syntax, Style and Grammatical Norms: English from 1500–2000*, Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt & Herbert Schendl (eds), 143–166. Bern: Peter Lang.

Rohdenburg, Günter. 2009a. Nominal complements. In *One Language, Two Grammars? Differences between British and American English*, Günter Rohdenburg & Julia Schlüter (eds), 194–211. Cambridge: CUP. doi:10.1017/CBO9780511551970.011

Rohdenburg, Günter. 2009b. Grammatical divergence between British and American English in the nineteenth and early twentieth centuries. In *Current Issues in Late Modern English*, Ingrid Tieken-Boon van Ostade & Wim van der Wurff (eds), 300–329. Bern: Peter Lang.

Rohdenburg, Günter. 2014. On the changing status of *that*-clauses. In *Late Modern English Syntax*, Marianne Hundt (ed.), 155–181. Cambridge: CUP. doi:10.1017/CBO9781139507226.013

Rohdenburg, Günter. 2017. Formal asymmetries between active and passive clauses in Modern English: The avoidance of preposition stranding with verbs featuring omissible prepositions. *Anglia* 135. 700–744.

Rudanko, Juhani. 2015. *Linking Form and Meaning: Studies on Selected Control Patterns in Recent English*. Houndmills: Palgrave Macmillan.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York NY: Harcourt Brace.

Visser, Fredericus Theodorus. 1963. *An Historical Syntax of the English Language, Part one: Syntactical Units with One Verb*. Leiden: Brill.

Vosberg, Uwe. 2006. *Die Große Komplementverschiebung. Außersemantische Einflüsse auf die Entwicklung satzwertiger Ergänzungen im Neuenglischen*. Tübingen: Narr.

Yañez-Bouza, Nuria. 2014. *Grammar, Rhetoric and Usage in English: Preposition Placement 1500–1900*. Cambridge: CUP.

Electronic sources

BNC = British National Corpus 1995. Version 1.0. BNC Consortium/Oxford University Computing Services. (100,000,000 words)

EAF = Early American Fiction 2000. Chadwyck-Healey. (34,634,660 words)

ECF = Eighteenth-Century Fiction 1996. Chadwyck-Healey. (9,702,696 words)

EEPF = Early English Prose Fiction 1997–2000. Chadwyck-Healey. In association with the Salzburg Centre for Research on the English Novel SCREEN. (9,562,865 words)

EPD = English Prose Drama 1996–1997. Chadwyck-Healey. (26,454,639 words)

ETC = Early Twentieth Century Corpus – a selection of British and American writings by authors born between 1870 and 1894. Source: Project Gutenberg. Compiled in the Research Project "Determinants of Grammatical Variation in English", University of Paderborn. (16,351,681 words)

ETC/B = British writings in the *ETC*. (4,801,408 words)

LNC = Late-Nineteenth-Century Corpus – a selection of British and American writings (complementary to the *EAF* and *NCF*) by authors born between 1830 and 1869. Source: Project Gutenberg. Compiled in the Research Project "Determinants of Grammatical Variation in English", University of Paderborn.

LNC/B = British writings in the LNC. (20,817,802 words)

MNC = Mid-Nineteenth-Century Corpus – a selection of British and American writings (complementary to the *EAF* and the *NCF*) by authors born between 1803 and 1829. Source: Project Gutenberg. Compiled in the Research Project "Determinants of Grammatical Variation in English", University of Paderborn.

MNC/B = British writings in the *MNC*. (10,082,876 words)

NCF = Nineteenth-Century Fiction 1999–2000. Chadwyck-Healey. (37,589,837 words)

NCF1 = First part of the *NCF* containing only those authors born in the eighteenth century (*1728–*1799).

NCF2 = Second part of the *NCF* containing only those authors born in the nineteenth century (*1800–*1869).

OED = The Oxford English Dictionary, 2nd edn on CD-ROM 1992 (Version 1.10). Edited by John A. Simpson & Edmund S. C. Weiner. Oxford: OUP.

t90, t92, t94 = *The Times & The Sunday Times* on CD-ROM 1990, 1992, 1994. Chadwyck-Healey/ProQuest. (119,476,791 words)

wridom1 = fictional component of the BNC (18,863,529 words)

CHAPTER 4

Finite causative complements in Middle English

Brian Lowrey

Université de Picardie Jules Verne, Amiens

This paper examines the use of causative verbs with two kinds of finite comple-
ments in Middle English (ME): a simple finite clause (V+*that*) and one involving
both a finite clause and a separate NP (V+NP+*that*). I attempt to determine in
which contexts these structures appear, and to what extent they are related. I
suggest that V+*that* is probably a survivor from Old English and V+NP+*that* a
later innovation. I also examine the factors likely to have contributed to the per-
sistence of V+*that* in ME, and to the emergence of V+NP+*that*, such as language
contact and the influence of other existing Middle English constructions which
may have played a role in the spread of V+NP+*that*.

Keywords: causatives, finite complements, separate NP, argument structure,
agentivity

1. Introduction

In this paper, I should like to examine a marginal but nonetheless intriguing feature
of Middle English (henceforth ME) complementation: the possibility for implicative
causative verbs to occur with finite complements. I borrow the term "implicative"
from Karttunen (1971) to refer to those verbs which entail that the action described
by their complement is indeed carried out. Consider, for example, the Present-day
English (henceforth PDE) causatives *make* and *have*, as exemplified in *She made/
had him open the door*. With *make* as well as *have*, the door is presumed to have been
opened. It is not normally possible to deny this, as there is an apparent contradiction
between the first and second clause of the sentence *She made/had him open the
door, but he didn't open it*. This property separates the implicative causatives from
non-implicative, manipulative verbs such as *order* or *command*. While verbs of the
latter type can be used to imply that the complement action has been performed,
this is a pragmatic inference and as such can be canceled: *She ordered/commanded
him to open the door, but he didn't open it*. This time, no contradiction occurs.

Where ME is concerned, the implicative causatives include *make*, *do*, and *cause*. Unlike their PDE equivalents, all of these still occur in ME with finite complements. More exactly, they seem to take two types of finite complement, illustrated by (1) and (2) respectively:

(1)   *The clennesse and the fastynge of us freres | **Maketh** that crist accepteth oure preyeres.*
'The purity and fasting of us monks causes Christ to hear our prayers.'[1]

(Chaucer, *Summoner's Tale*, 1883–1884)

(2)   *Myn heed is toty of my swynk to-nyght, | That **makes** me that I ga nat aright.*
'My head is dizzy from my night's work, which causes me to lose my way.'

(Chaucer, *Reeves Tale*, 4253–4254)

In (1), the complement of causative *make* is the finite clause introduced by *that*. In (2), on the other hand, the causative appears to have two complements, both the finite clause and an additional NP which is co-referential with the subject of the clause, but apparently not part of the clause. I shall call these structures V+*that* and V+NP+*that*, respectively. That such constructions should still be found in ME is intriguing, given that while finite complements were relatively common with causatives in Old English (henceforth OE), the shift from finite to non-finite complementation remains one of the elements that characterize the transition from the OE to the ME period (see, e.g., Manabe 1989; Los 2005). I shall try to determine here to what extent these constructions can be analyzed as "relics" from an earlier period, or whether they are ME innovations. I shall examine the contexts in which they tend to appear, and the semantic values associated with them. I shall also try to ascertain whether V+*that* and V+NP+*that* are separate constructions or variants of a single basic model, and ask to what extent examples such as (1) and (2) suggest that causatives such as *make* were two- or three-place verbs, in terms of argument structure.

2.   The frequency of finite causative complements

I should perhaps begin by confirming the essentially marginal status of finite causative complements in ME. To do so, I quote here the figures given in Manabe (1989), who compares the relative distribution of finite and non-finite complements across a range of ME texts. Manabe's tables summarize occurrences of *maken* (1989: 116) and *don* (1989: 105):

---

1.   All the translations and glosses given here, unless otherwise stated, are those of the author, who obviously accepts full responsibility for any errors or omissions.

maken

	EME	14th c.	15th c.	Total
Finite complement clauses	9	6	0	**15**
Infinitive complement clauses	177	178	143	**498**
+S	162	168	128	
−S	15	10	15	

don

	EME	14th c.	15th c.	Total
Finite complement clauses	11	1	0	**12**
Infinitive complement clauses	106	88	32	**226**
+S	103	64	25	
−S	3	24	7	

The term "+S" denotes the traditional AcI ("*accusativus cum infinitivo*") construction, whereas "−S" designates the type of infinitive complement, common with causatives and perception verbs in OE, in which the subject of the infinitive is left unexpressed.[2] This is the construction dubbed "V+I" by Denison (1993: 165). Manabe is chiefly concerned with infinitives, and does not make a distinction between V+*that* and V+NP+*that*. Nonetheless, his figures show that finite complements in general occur only occasionally with causatives throughout the ME period, suggesting that, like *that* complements to other verbs, they become less frequent over time and have died out by the 15th century. The very small numbers of finite examples make it difficult, though, to draw reliable conclusions.

My own findings largely confirm those of Manabe. I ran a search through a sample of ME prose and poetic texts, consisting of over 1,400,000 words. Most of the texts contained no instances of the target structures. This was the case for: Laʒamon's *Brut* (ll. 9,000–15,000), *Hali Meiðhad*, *Sawles Warde*, *Ancrene Wisse* (Chapter 3–8), *King Horn*, *Dame Siriþ*, *The Peterborough Chronicle*, *The History of the Holy Rood Tree*,

---

2. To illustrate Manabe's construction types, here are two examples involving *do*.

(i) *For David in hise dayes dubbed knyghtes, / And **dide** hem sweren on hir swerd to serven truthe evere.*
'For David, in his day, dubbed knights, and had them swear upon their swords ever to serve the truth.' (*Piers Plowman*, 1.98–99)

(ii) *And I shal covere youre kirk, youre cloistre **do** maken / Wowes **do** whiten and wyndowes glazen.*
'And I shall cover your church, have your cloister built, have the walls whitened and the windows glazed.' (*Piers Plowman*, 3.60–61)

Example (i), where *hem* is the subject of the infinitive *sweren*, would be an example of the "+S" construction. Example (ii), in which the subject of the infinitives *maken*, *whiten*, and *glazen* is understood, is an instance of what Manabe calls the "−S" construction.

Athelston, *Cleanness*, *Patience*, the *Pearl*, *Sir Gawain & the Green Knight*, *Sir Orfeo,*
The Siege of Jerusalem, *Mandeville's Travels*, *Octovian*, the *York Mystery Plays*, the
works of Richard Rolle in Ogilvie-Thomson's *Richard Rolle, Prose and Verse*, and *The*
Prose Rule of St. Benet. Finite causative complements were found in *Vices & Virtues*,
Havelok, *Owl & Nightingale*, Langland's *Piers Plowman*, Gower's *Confessio Amantis*,
Chaucer's *Troilus & Criseyde*, and the *Canterbury Tales*, as well as in Malory's *Morte*
d'Arthur, the *Paston Letters*, and a corpus of prose texts by Chaucer.[3] A summary of
the results for those texts in which the target structures occur is given in Table 1.[4]
The table, corresponding to Manabe's model, has been expanded to include the other
principal ME causatives, *gar*, *let*, and, in later ME, *cause*:

Table 1. Finite and non-finite causative complements, texts containing V+*that* and/or
V+NP+*that*

	do	*make*	*gar*	*cause*	*let*
Finite complement clauses[*]	16	26	0	6	0
+*that*	12	17	0	2	0
+NP+*that*	4	9	0	4	0
Infinitive complement clauses	367	460	22	151	216
+S	192	436	14	146	0
−S	175	24	8	5	216

* The figures reflect the number of complement clauses: two or more clauses coordinated under the same matrix verb are counted as two or more separate occurrences.

The comparative scarcity of finite causative complements in ME contrasts with the
situation in OE (at least the West Saxon literary variety), as can be seen from the
corresponding figures for *hatan* and *gedon*, the two most common OE causatives,
in a late OE text, the *Heptateuch* (see Lowrey 2012 for more detailed information).
The *Heptateuch* alone (approximately 75,000 words) yielded 15 finite complement
clauses with the verb *(ge)don*:

Table 2. Finite and non-finite causative complements in the *Heptateuch*

	hatan	*gedon*
Finite complement clauses	1	15
+*that*	0	15
+NP+*that*	1	0
Infinitive complement clauses	45	0
+S	*16*	*0*
−S	*29*	*0*

---

3.   Comprising the *Treatise on the Astrolabe*, *Boece*, the *Tale of Melibee*, and the *Parson's Tale*).

4.   Figures for each individual text are given in Appendix 1.

While *hatan* almost exclusively selects infinitive complements, causative *(ge)don* is found in this text with finite complements only. Before examining the distribution of V+*that* and V+NP+*that* in ME, it might therefore be of interest to determine under what conditions finite structures appear in OE.

3.   Finite causative complements in OE

3.1   "First" and "second" causatives

Causative constructions in English have always been sensitive to a certain number of parameters, which determine to a greater or lesser extent which causative verb and which complement type will co-occur in a given context (Lowrey 2002). These parameters can be clustered around two "poles" situated at opposite ends of a scale, which I shall call "first" and "second" causatives, respectively:

First causative	*Second causative*
←	→
Agentive causer/causee	Non-agentive causer/causee
+specific, −generic	−specific, +generic
+concrete, −hypothetical	−concrete, +hypothetical
Complement clause denotes an action, often an "accomplishment" (Vendler 1967)	Complement clause denotes a state
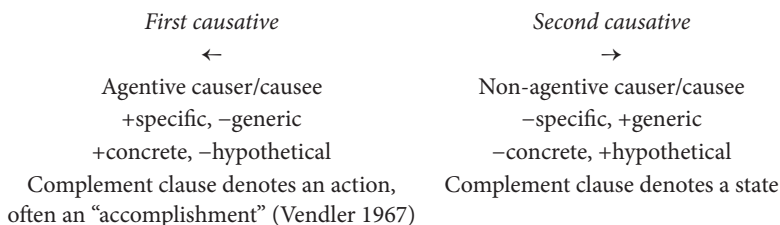
Figure 1. First and second causatives, OE and ME

Typical first causative situations involve volitional causers and/or causees, specific rather than general contexts, concrete rather than hypothetical situations, and caused events rather than states. Interestingly, the same factors prove to be at work both in OE (Lowrey 2010, 2012) and ME (Lowrey 2002, 2013a).

A first causative situation in OE is typically marked by the verb *hatan* with an infinitive complement, as in (3), an instance of Manabe's "−S" construction:

(3)   *Ic þe **het** feccan þæt þu mine fynd wirigdest and þu nu þriwa hig bletsodest.*
      'I had you brought here so that you would curse my enemies, and you now
      have blessed them three times.'   (*Heptateuch*, 152: 3)

Hatan (literally 'order, command') functions here, as it often does in OE (Royster 1918; Lowrey 2013b), as a full-fledged implicative causative, hence the use of *had* in the modern English translation. The causer (*Ic*) is an Agent, as is the (implicit) causee (*þe*), which are typical first causative features.

In fact, the majority of causers in the *Heptateuch* (as in most surviving OE texts) are volitional agents. However, when a non-agentive causee appears, we tend to find the finite complement construction shown in (4), with the verb *(ge)don*:

(4)   *ic **gedo** þæt eow bið ægþer heard, ge heofene ge eorþe.*
      'I shall make them both hard for you, heaven as well as earth.'

<div align="right">(*Heptateuch*, 136: 24)</div>

The presence of an inanimate causee, the stative complement and the future context are all second causative features. It is in this kind of environment with *(ge)don*, the typical second causative verb, that V+*þæt* commonly occurs in OE. As Table 2 shows, *(ge)don* does not appear with an infinitive complement in the *Heptateuch*, and all but two of the 15 instances where the verb is used in the V+*þæt* construction involve a non-agentive causee (see Lowrey 2012: Section 3.3).

3.2   V+NP+*þæt* in OE

At the same time, OE also had a V+NP+*þæt* complement construction, which was especially common with non-implicative, manipulative verbs such as *bebeodan* and *biddan* (see the figures in Lowrey 2012: Section 3.3):

(5)   *Of þæs treowes wæstme þe stent omiddan neorxenawange, God **bebead** us þæt we ne æton, ne we þæt treow ne hrepodon, þy læs þe we swelton.*
      'Of the fruit of the tree that stands amid the Garden, God ordered us that we should not eat, nor touch the tree, lest we die.'   (*Heptateuch*, 12: 13)

Both the NP *us* and the finite clause, apparently, are independent arguments of *bebeodan*. The verb has a three-place argument structure which seems to match its manipulative semantics quite neatly. V+NP+*þæt*, however, is restricted in OE to manipulative verbs of this type. Significantly, causative *(ge)don*, when used with a finite clause, only appears in V+*þæt*, never in V+NP+*þæt*.

4.   Finite causative complements in ME

In the light of the elements presented in Section 3, we can now look more closely at the distribution of V+*that* and V+ NP+*that* in ME.

4.1   V+*that* in ME

With regard to the diatopical distribution of ME causatives, Mustanoja (1960: 601) observes that "the causative *do* seems to be favored in the eastern parts of the country, while *make* and *let* seem to prevail in the West. The causative verb preferred in the more northern areas is *gar*". Although this is essentially correct, it remains

something of a simplification. It might be more accurate to say that *gar*, *let*, and *do* are the commonest first causatives in their respective dialects. The commonest second causative, on the other hand, is *make*, in all dialects. It is with the verb *make*, in typically second causative contexts, that finite V+*that* complements appear most often in ME. This is the case in (1) above, and also in (6) and (7):

(6)   *Bot Jelousie of his untrist / **Makth** that full many an harm arist, / Which elles scholde noght arise.*
      'But Jealousy, through his untrustworthiness, makes many a wrong to occur, that otherwise should not occur.'   (Gower, *Confessio Amantis* V.717–719)

(7)   *But thilke ordre, (…) **makith** that the causes rennen and assemblen togidre.*
      'But that same order makes the causes run and come together.'
                                                        (Chaucer, *Boece*, Liber V Prosa 1, 92)

As the above examples show, finite complement clauses are found both in prose and in poetry. In (6) and (7), the causer is non-agentive, and the situation is generic in character. Once again, these are typical second causative features.

V+*that* is sometimes found with an agentive causer, usually with the verb *don*:[5]

(8)   *and swa ðu miht **don** ðat all ðat euel ðe mai wanden te michele gode*
      'and so you can make it that all that evil may turn to great good for you'
                                                        (*Vices and Virtues*, 29: 13)

Notice, however, the presence of other second causative features here. The causee is inanimate, while the sentence has a hypothetical value. Sentence (8) illustrates a further property of the finite construction, namely that it enables a modal to appear in the complement clause.

Overall, the V+*that* examples I have found correspond to second causative situations, with non-agentive causers and/or causees and stative complements. None of my examples appear in typical first causative contexts.[6] V+*that*, in other words, seems to occur in ME in much the same environments as in OE. It would therefore be possible to see the ME construction as a kind of relic, which dies out over time.

It seems to die a very slow death, however, if this is indeed the case. As (9) below shows, causative V+*that* has not yet become completely obsolete by the 15th century, in contrast to what Manabe's (1989) tables seem to suggest (see Section 2). In fact, the pattern remains at least marginally productive even in late ME:

---

5.   The causer is agentive in 11 instances out of 27 in the V+*that*-construction (8 with the verb *don*, 3 with *maken*), and in 2 instances out of 17 in V+NP+*that* (both with the verb *don*).

6.   See the full list of examples given in Appendix 2.

(9) *But I conceyve now the trought, and þat sikenesse* **caused** *that she be absent here.*
'But I now conceive the truth, and that sickness caused her to be absent here.'
<div align="right">(*Paston Letters*, 369: 8)</div>

Example (9) is taken from a letter written by Margaret Paston, dated 1472, and which presumably constitutes an authentic example of the epistolary English of the time. Of particular interest here is the fact that V+*that* occurs with the verb *cause*. The use of *cause* as a periphrastic causative is a ME innovation. Much like *make* before it, *cause* enters the causative group at the non-agentive end of the scale given in Section 3.1 (see Lowrey 2013a). The finite complement construction is therefore still productive enough to be used with a new lexical item at this stage, associated with the same kind of second causative semantics as *(ge)don* in OE.

4.2 Survival of V+*that*

It is not altogether clear, at a stage in the history of English where finite complements of causative and manipulative verbs generally give way to non-finite structures, why causative V+*that* should remain at least partially productive throughout the ME period. As the modals grammaticalized and lost their infinitive forms, the persistence of the finite construction did at least allow modals to be incorporated in the complements of causative verbs, which may have contributed to the survival of V+*that*. It is unlikely to have been more than a contributory factor, however. Many of my V+*that* examples do not contain a modal, and other languages, including modern English, function quite happily without an equivalent causative + modal complement construction.

Another explanation might be the influence of Anglo-Norman.[7] Anglo-Norman French, much like modern French, possessed a construction in which causative *faire* appeared with a finite *que*-complement, as in the following example:

(10) *Ne lerra   qe Daire   a destresces ne   assaille;/ Si par [alcun]*
      NEG will.allow that Darius forcibly   NEG attacks if by any
      *engin puit **fere** qu' il s'en aille.*
      device can make that he go.away.SUBJ
      '[Alexander] will not permit that Darius attacks with force if by some ruse [he] can get him to go away.'
<div align="right">(Thomas of Kent, *Le Roman de toute Chevalerie*, 1858–59)</div>

---

7. I would like to thank Richard Ingham for his comments and suggestions regarding the influence of Anglo-Norman.

As Ingham (2012) points out, Anglo-Norman was probably more extensively used in medieval Britain than has hitherto been assumed. He quotes Short (2007), who suggests that even as late as the early 15th century some 20% of the population would have been able to speak both French and English. The degree of bilingualism would presumably not have been high enough to cause English complementation patterns to be aligned with those of French, particularly at a time when the general trend was for infinitive complements to replace finite ones. It may, on the other hand, have been sufficient to prolong the survival of what would presumably have been felt to be a somewhat archaic form. Interestingly, John Gower who, as we have seen, used V+*that* in his English poetry also used causative *faire* with a finite *que-* complement in his Anglo-Norman French works:

(11) *Trop est ce   vice malsené, Que* **fait   que** *Dieus est coroucé.*
      too is   this sin unwise   that makes that God   is   angry
      'This sin is too unwise, that makes God angry.'

                                        (*ANHT*, Gower, *Mi'rour de l'Omme*, 4006)

4.3   V+NP+*that*

Let us now turn to the V+NP+*that* construction. I shall examine, in this section, the nature and the distribution of V+NP+*that* in ME, and consider some of the factors which may have led to its emergence.

4.3.1   *Distribution of V+NP+*that

The first point to make here is that V+NP+*that* seems to occur in very much the same type of second causative environments as the "simple" finite complement. Some examples are given below:

(12) *Mighte no lesynges, ne lyere, ne los of worldly catel. /* **Maken** *hym, for any mournynge, that he nas murie in soule.*
      'No deceit, no liars, no loss of worldly goods could make him, however sad the situation, not be merry in his soul.'   (Langland, *Piers Plowman*, XIX: 294)

(13) *For which, where as his peple therbifore / Hadde loved hym wel, the sclaundre of his diffame /* **Made** *hem that they hym hatede therfore.*
      'For which reason, whereas his people had previously loved him well, the scandal of his bad reputation made them hate him for it.'
                                              (Chaucer, *Clerk's Tale*, 729)

(14) *Ðies* **dieð** *ðe manne ðe ðurh hire is beswiken ðat he twifealdeð his senne.*
      'This causes the man that is deceived by it to double his sin.'
                                              (*Vices and Virtues*, 5: 20)

(15)   *They bothe told me þat ye schold veryly a ben at home be-for Crystmas, and that **causyd** me that I wrot not to yow non answer.*
'They both told me that you should indeed have been home before Christmas, and that caused me not to write you an answer.'   (*Paston Letters*, 276: 10)

(16)   *'Gramercy,' seyde sir Gareth, 'myne owne lady. For this rynge is passynge mete for me; for hit wille turne all maner of lykenesse that I am in, and that shall **cause** me that I shall nat be knowyn.'*
'"Thank you very much, my own lady," said sir Gareth, "for this ring is just right for me, for it will transform whatever appearance I have and that shall cause me not to be known."'   (Malory, *Morte d'Arthur*, 214: 4)

As the above examples show, V+NP+*that* is found, just like V+*that*, in both prose and poetry, in late (*Morte d'Arthur*) as well as in early (*Vices & Virtues*) ME texts. It, too, occurs frequently with *make*, and also with *don* and *cause*.

In each of the above examples, the causer is non-agentive, as in most cases is the causee too. The complement often evokes some kind of state or a property of the causee. I have thus far found only one apparent counter-example to this tendency, in *Havelok*:

(17)   *Hwan he hauede **don** him for drede, / Þat he sholde hire spusen, and fede, / And þat she sholde til him holde, / Þer weren penies þicke tolde.*
'When he had made him, for fear, agree to marry and feed her, and her agree to hold to him, many pennies were counted out.'   (*Havelok*, 1169–1172)

As the use of the term *for drede* suggests, this example is apparently situated towards the agentive, manipulative end of the causative scale, under the interpretation suggested by my translation. Note that even here one finds a modal in the complement, a factor which may partially explain the recourse to a finite structure. If this interpretation is the correct one, (17) would seem to present an interesting case of co-ordination of the two types of finite causative complement.[8]

On the other hand, (17) could also, in the context of the poem, lend itself to a different interpretation, if *hauede do* is taken to mean something like 'put himself to it that…' 'made himself accept that…' (*MED* s.v. *don* (v) 6(a)). The referent of *he* would, in this case, be Havelok himself, rather than the evil Godrich, who seeks to force Havelok to marry Goldborow, the referent of *she*.[9] Since both interpretations are compatible with the context, and given the otherwise rather exceptional

---

8.   While the first complement clause in *don him (…) þat he sholde hire spusen* looks very much like a V+NP+*that* structure, this does not seem to be the case for the second, *þat she sholde til him holde*, whose subject, *she*, is obviously not coreferential with the extra-clausal NP.

9.   I am indebted to one of the anonymous reviewers for pointing this out to me.

semantics of V+NP+*that* here, the second interpretation may well prove to be the correct one.

With the possible exception of (17), then, the examples of V+NP+*that* that I have found seem to share to a very large extent the semantic properties associated with the simple finite construction.

4.3.2   V+NP+that *and the evolution of causative* make

Not only does V+NP+*that* survive until the very end of the ME period, there is also some evidence that it remains at least marginally productive in Early Modern English (EModE) with causative *make*. Denison (1993: 181) cites two EModE instances of the construction, from Pepys' diary, one of which is reproduced below:

(18)   *Only want of practice* **makes** *her she cannot go through a whole tune readily*.
                    (Pepys, *Diary* VIII 221.18, example quoted by Denison 1993: 181)

In the light of the history of causative *make* proposed by Terasawa (1985), Denison (1993: 207) raises the question as to whether such examples might be a sign that *make* has acquired a three-place argument structure. To summarize Terasawa's hypothesis, *make* functions initially as what he calls a "pure" (i.e. two-place) causative, with "resultative" (i.e. implicative) meaning. Later, it comes to acquire an additional function as an "agentive" causative, with both resultative and "act on" (i.e. manipulative) meanings, to use Terasawa's terminology. In other words, *make* becomes a three-place, manipulative verb, not unlike *bebeodan* in OE, but with implicative meaning.

Terasawa is correct in asserting that *make* is increasingly used as a first causative. Indeed, *make* enters the causative group somewhere towards the right-hand end of our causative scale (see Section 3.1, Figure 1) and subsequently moves leftwards (Lowrey 2013a). However, I suspect, for a number of reasons, that neither my ME examples, such as (12), nor Denison's EModE examples, are connected with this change.

First of all, the semantics simply do not fit. As we have seen, V+NP+*that* is typically associated with second causative semantics, which correspond not to Terasawa's agentive causative but rather to his resultative, two-place causative. If V+NP+*that* really were the result of *make* becoming more manipulative in character, then one would expect to find it typically in cases like (17). And yet (17) – if the agentive interpretation is correct – remains something of an exception, the only instance of its kind that I have discovered so far.

Secondly, among the ME periphrastic causatives, V+NP+*that* complements are not restricted to the sole verb *make*. As (14), (15), and (16) show, they were also used with *don* and even with *cause*, which is not used as an agentive causative at all in ME (although it does briefly acquire typical first causative functions in EModE;

see Lowrey 2002). The possibility of occurring in the V+NP+*that* construction seems to have been a more general property of periphrastic causatives, rather than a phenomenon linked to the specific development of one verb.

Finally, it should be pointed out that the extension of *make* to first causative uses is indeed reflected by a change in complementation patterns. However, the relevant change here is the tendency for bare infinitive complements, Visser's (1973) VOSI construction, to occur in agentive contexts and for *(for) to*-infinitive complements (or VOSI[to]) to appear with non-agentive causers and causees. This is a clear pattern which emerges in ME/EModE (Lowrey 2002, 2013a). V+NP+*that* causative complements are found both in early ME and in late ME second causative contexts, and seem to have little to do with the extension of *make* to more agentive contexts.

4.3.3   CLAN sentences

If, then, we reject the idea that V+NP+*that* is some form of nascent three-place manipulative structure, how are we to explain its existence? At first sight, at least, V+NP+*that* appears to resemble what Warner (1982: 91–9) calls a "CLAN sentence" (for CLAuse and Nominal). An example of a CLAN sentence, borrowed from Warner, is given below:

(19) *Herodius aspiede Joon many gatis, how he myȝte be do to deþ.*
       'Herod observed John in many ways, how he could be put to death.'
                                    (*Wyclifite Sermons*, quoted by Warner 1982: 95)

In a CLAN sentence, an NP (*Joon*, in this instance) and a finite complement clause stand in some kind of appositional relationship, or, as Denison (1993: 206) puts it, together they occupy "a single position in the argument structure of the higher verb". Warner himself suggests that CLAN sentences share the following basic property:

> If we factor any CLAN sentence into 2 sentences, one containing the nominal and the other containing the clause, then we can parallel the construction with the matrix expression in each case, and we always find that the resulting two statements are appropriate and true in context.   (1982: 95)

In other words, (19) entails both (i) and (ii):

i.   *Herodius aspiede Joon.*
ii. *Herodius aspiede how Joon myȝte be do to deþ.*

Warner's factoring operation, however, turns out to be problematic with causative constructions such as my example (16), partially reproduced here:

(16)   …*that shall cause me that I shall not be knowyn*

While this sentence quite clearly entails *that shall cause that I shall not be knowyn*, it obviously does not entail *that shall cause me*. The independent NP cannot function, alone, as an argument of the causative. Presumably, therefore, (16) is not a CLAN sentence, and apparently not a three-place structure either.

4.3.4   *A comparison with perception verbs*

It might be of interest, in order to have a better understanding of causative V+NP+ *that*, to take a look at complementation patterns occuring with perception verbs in ME. It has often been pointed out that perception verbs and causatives seem to share a number of properties (see, for instance, Quirk et al. 1985: Section 16.52). Of particular relevance to our study is the fact that Kopytko (1985) finds cases in ME of V+NP+*that* constructions occuring with verbs of sensory perception that are apparently similar to those with causative verbs, as in the following example:

(20)   *Þe knyghtis of Rome saw Vaspasyan at he was a noble man.*
       'The knights of Rome saw that Vespasian was a noble man.'
                           (*Alphabet of Tales*, 427, example quoted by Kopytko 1985: 27)

Kopytko takes the existence of examples such as (18) as evidence that verbs such as *see* or *hear* were in fact three-place verbs in ME. In a reply to Kopytko, Fischer (1987) disagrees, arguing convincingly that perception verbs in ME were two-place verbs, just like their present-day counterparts. Instead, she suggests that the separate NP in structures such as (20) is in fact topicalized to the front of the complement clause, without at any stage becoming an argument of the higher verb. Using the framework available at the time, she suggests the structure V [$_{S'}$ *TOPIC* [$_S$ *NP VP*]]. As she puts it: "the NP is a part of the S′ clause, and has been moved out of the clause into TOPIC position, leaving a resumptive pronoun behind" (1987: 63).

An analysis of this type would, in semantic terms, seem to fit at least some in-stances of causative V+NP+*that*. In many of the examples that I have found, some kind of prominence is given to the role of the causee, in as far as (s)he is the – often human – entity that undergoes a change of state or is affected by the action of an of-ten inanimate or abstract causer, although not at all manipulated in the traditional, agentive sense of the term.

As we saw in the previous section, the NP in question seems not to be an argu-ment of the causative in its own right, but rather to belong in the embedded clause. Should this analysis prove to be correct, another property common to both caus-atives and perception verbs comes to light: the possibility for both to select finite complements with a topicalized NP. And of course, the fact that such complements could occur with causatives would then tend to confirm that Fischer's analysis of perception verb V+NP+*that* is probably correct.

Concerning causatives, the topicalization analysis would also explain why the semantics and the distribution of V+*that* and V+NP+*that* are so similar. The latter construction, seen from this perspective, is simply a variant of the former, one which emerges during the ME period. This in turn would explain why the distributional and semantic properties of ME V+NP+*that* are so different from those of the OE construction, which, as we saw in 3.2, appeared essentially with non-implicative, manipulative, agentive verbs. ME V+NP+*that* is not a continuation of the OE construction.

However, while a topicalization analysis would seem to fit examples such as (13), (14), (15), or (16), it cannot be applied straightforwardly in all cases. Of the 14 occurrences of V+NP+*that* that I have found, two appear problematic. In (12) and (17), the NP denoting the causee precedes the adverbials *for any mournynge* and *for drede*, respectively, which apparently form part of the higher clause,[10] suggesting that the NP, rather than being located at the front of the embedded clause, is also a part of the matrix clause in these instances.[11]

4.3.5   *Links to other causative constructions in ME*

In another work, Fischer (1990) suggests an alternative analysis for the emergence of V+NP+*that* with causative *make*, quoting the following example:

(21)   *the wilde loves rage, / In which no man him can governe, / Hath **mad** him that he can noght werne*
       'the wild passion of love, in which no man can control himself, has made him unable to refuse'   (Gower, *Confessio Amantis*, 1 2620–2623)

Fischer suggests that the finite complements in examples of this type "could and presumably ought to be interpreted as consecutive clauses" (1990: 248). The suggestion is made only briefly, with no further explanation given, but Fischer presumably

---

10. Unless the adverbial PPs in question are in fact parenthetical structures of some kind, which could have been inserted after the NP without altering the basic structure of the sentence. In this case the NP could still be part of the complement clause. I am grateful to Hubert Cuyckens for pointing this out.

11. Kroch and Taylor (1997) mention another construction in which a constituent of the embedded clause can be "preposed to the immediate left of C^0" in ME, which they call the "doubly filled comp sentence" (1997: 315). However, this would not seem to fit our examples (12) and (17). Doubly filled comp, according to Kroch and Taylor, can be ascribed to Scandinavian influence, and as such is restricted to Northern texts in ME, absent from texts of Midland and of Southern origin. While *Havelok*, a North-East Midland text, from which (17) is taken, displays a number of Northern features (the presence of causative *gar*, for example), and might be considered to reflect considerable Scandinavian influence, this is not the case of *Piers Plowman*, the source of (12). Notice, too, that in both (12) and (17) the relevant NP does not immediately precede the complementizer *that*.

interprets (21) as meaning that the passion of love changed or at least did something to the referent of *him*, with the result that he could no longer refuse. A possible problem for this hypothesis is the fact that it would require the NP, *him* in (21), to function as an argument of *make* in its own right. However, *make* is not attested in ME with a NP complement, standing alone and referring to a human entity, with a 'changed, affected' meaning. The construction *is* found with the meaning 'bring about, create' as in (22):

(22) *Ihesu lord, þat **madyst** me. Forȝeue me*
      'Jesus, lord, that made me. Forgive me'
                    (MED s.v. *maken* 1(a) <u>c1460 *Iesu þat art heuene*</u>)

Make here, though, obviously does not have the same meaning as in (21). If the verb does indeed mean something like 'changed' or 'affected' in (21), the simple V+NP construction is unlikely to have been the source of V+NP+*that*. This would seem to rule out another possibility mentioned by Fischer, in which the *that*-clause "does not constitute another object" next to the NP but functions rather as a kind of explanatory clause. In Fischer's words, it "explains the nature" of the NP (1990: 99). However, this again would require that the NP could stand alone, with the relevant meaning, as an argument of *make*, which is not the case in ME.

    On the other hand, some evidence that *make* could indeed express the 'affected' meaning in the NP+finite complement construction is provided by a second, somewhat unusual, example quoted by Fischer (1990: 249):

(23) *Merlion did **make** King Arthur that Sir Gawain was sworn to tell of his adventure and how he would give no mercy unto the knight.*
      'Merlin made king Arthur act with the effect that sir Gawain was sworn to tell of his adventure and how he would show the knight no mercy.'
                          (Malory, *Morte d'Arthur*, 67: 30–31)

This, as far as I can ascertain, is the only causative example of its type, in which the independent NP and the subject of the finite clause are not coreferential. Fischer herself calls (23) a "dubious case" (1990: 248), noting that "in the C text the verb *make* has been replaced by *desire*, with which a *that* clause is quite regular" (1990: 249).[12] Example (23) nonetheless seems to imply that Merlin did something to Arthur, with the result that Gawain pledged to relate his adventure, and remains consistent with the "consecutive clause" analysis.

---

12. The expression "C text" here refers to the version of Malory's text published by Caxton, reproduced in H. Oskar Summer's 1899 edition of the *Morte d'Arthur*. Caxton is believed to have modified parts of the text. The example quoted by Fischer, as well as my example (16), is taken from Vinaver's edition of the (presumably earlier) Winchester manuscript.

Perhaps a more plausible explanation of the emergence of V+NP+*that* structures involving *make* lies in the spread of the verb to the so-called "small clause" construction, Denison's (1993: 183) "V+NP+Pred" in which "Pred" denotes a "predicative adjunct", as in:

(24)   *oferlifa on hete and on wete* **macað** *þene mon un-halne*
       'overliving in eating and in drinking makes the man unwhole'
       (*Lambeth Homilies* 101: 11)

From a semantic point of view, the construction in (24) is closer to the V+NP+*that* of (21) in that, here too, something is done to the referent of the NP *þene mon*, who undergoes a change of state as a result. *Make(n)* takes over from *do(n)* in causative V+NP+Pred between late OE and early ME, around the time at which V+NP+*that* appears, as part of a series of more general changes affecting the causative group at the time. Simplifying matters somewhat, causative *hatan* is lost, replaced in first causative contexts (and in the V+I construction) by *do* in Eastern dialects, by *gar* in the North, and by *let* in the South and West.[13] *Make*, generally, is relatively rare in OE (Royster 1918; Bock 1931; Ikegami 1981), but rapidly becomes a common second causative in all ME dialects, shifting increasingly to first causative contexts during the ME period (Lowrey 2013a). The earliest causative occurrences of *macian* are found in late OE, notably in the writings of Ælfric. *Macian* occurs in at least two types of V+NP+Pred construction, illustrated by (25) and (26) respectively. In the second type, the "Pred" element is a PP introduced by *to*:

(25)   *Swaðeah Paulus siððan forestop Stephanum on Godes gelaðunge mid menigfe-*
       aldum geswincum, þone ðe he ær ehtende martyr **gemacode**.
       'Nevertheless Paul, through many labours, afterwards, in God's congregation,
       preceded Stephen, whom he had previously made a martyr through persecu-
       tion.'   (*ÆCHom 2*: 82: 22–24)

(26)   *Se forma heafodleahtor is on leden* superbia *and on englisc modygniss. Seo*
       macode *to deoflum þa wlitigan encglas, þe wunedon on heofonum.*
       'The first capital sin is called superbia in Latin, and pride in English. It made
       devils of the beautiful angels that lived in heaven.'
       (*HC, Aelfric's letter to Wulfstan* 401–404)

---

13. This is a very broad simplification. In fact, the situation is more complex in the South and West, where *let* is used with a causative sense only in the V+I construction; see Lowrey (2002, 2013a, 2013b) for a more detailed presentation of these changes.

The Helsinki Corpus contains no instances of *macian* V+NP+Pred prior to the O3[14] period, and just two clear O3 examples, including (26).[15] In contrast, no fewer than 14 examples occur in the *Lambeth Homilies* alone, a sign that the construction was gaining momentum at the time V+NP+*that* appeared. *Make*+NP+*that* could therefore have been "primed" in the sense of Traugott & Trousdale (2013: 54) by two existing constructions: by V+*that*, with its essentially second causative semantics and finite clause, and also by the expanding *make*+NP+Pred construction, in which the NP stands in much the same relationship to *make*, also affected by the causative action and undergoing a change of state. In both of the existing constructions, *make* at this stage was supplanting *do*.

Of course, while such an explanation might lend itself quite readily to the rise of V+NP+*that* with *make*, one would still need to explain how verbs such as *cause* and *do* come to be used in this construction. The case of *cause* would presumably involve the extension of an existing construction to a new verb. As we have seen, *cause* is a relatively late addition to the causative group, as a new second causative. The earliest example with a simple nominal complement quoted by the *Oxford English Dictionary* (*OED*) is from 1340, while the earliest occurrence with an infinitive complement is from Gower. *Cause*+NP is restricted to the 'bring into being', 'give rise to' sense, and is not used to mean that the causer affected or modified the referent of the NP. And while *cause* is attested in the V+NP+Pred construction, such uses appear to have been infrequent in ME. Both the *OED* and the *Middle English Dictionary* (*MED*) quote just two examples (the same ones), in which the "Pred" element is a *to*-PP. One of these is reproduced below:

(27)   *What **causet** the kyng to his cleane yre, Tellis not the trety.*
　　　'What brought the king to pure wrath, the story does not tell'
　　　　　　　　　　　　　(*MED* s.v. *causen* 2(b), *Destruction of Troy* c1540[?a1400])

Cause+NP+Pred and *cause*+NP would therefore seem to be unlikely sources of *cause*+NP+*that*. The spread of a new verb to an existing construction appears more likely. *Cause*, the new second causative, would therefore have begun to be used in V+NP+*that* by analogy with *make*, its predecessor in that role. It may be significant that the later examples of V+NP+*that* in my corpus, in the *Paston Letters*, and in

---

14. O1 pre-850; O2 850–950; O3 950–1050; O4 1050–1150.

15. The second example occurs in *Passion of St. Margaret* 181–182. I have discounted examples such as *Hi **maciað** eall heom sylfum to woruldwlence and to idelre rence* ('They turn everything to their own worldly pride and vain pomp', Wulfstan, *Institute of Polity* 346–347), where *macian …to* seems to mean something like 'apply, turn… to/use… for', and not to have the "full" causative sense.

Malory's *Morte d'Arthur* involve *cause* and not *make*, suggesting perhaps that the former was taking over from the latter in this context, although it should be stressed that the number of examples is far too small to enable reliable judgments to be made.

(Ge)don, the OE second causative, on the other hand, is common in V+NP+Pred in OE from a much earlier date, found in O2 texts in the Helsinki Corpus:

(28) *he astereð þone rodor & þa tunglu, & þa eorðan **gedeð** stille*
      'It moves the firmament and the stars, and makes the earth be at rest.'

                                                                    (*HC, Boethius* 1137)

However, *do* is already losing out to *make* in this construction in early ME. If the rapid spread of *make* to V+NP+Pred is a probable factor in the rise of *make* in V+NP+*that*, the same explanation cannot be applied directly to the case of *do*. *Don*+NP+*that* is not attested at all in OE (Fischer 1990: 108). Only in early ME texts does the verb begin to appear in this construction, at about the same time as *make*. It is possible that an extension to *do* of the new V+NP+*that* structure, again by analogy with *make*, occurred in very late OE, particularly in dialects where causative *do* remained frequent, presumably as what Royster (1922: 347) calls a "folk use", only surfacing in written texts during early ME, although this would be difficult to prove.

There is a further semantic problem for the hypothesis that sees V+NP+*that* emerge by analogy with V+NP+Pred. While most of the complement clauses in the V+NP+ *that* examples that I have found designate a property of the referent of the NP or include a modal, describing what the referent of the NP could or might do (see Appendix 1), and would appear compatible with the semantics of the V+NP+Pred construction, some examples, such as (14) and (15), in which the clause describes an action, do not fit this pattern. One could perhaps argue that the clause in (14) lends itself to a near-stative reading: the referent of *ðe manne ðe ðurh hire is beswiken* performs no deliberate act, and the meaning is close to 'becomes twice as sinful', which would be a form of change of state. But the same reasoning cannot be applied to (15). Although the causer here is inanimate, and we are some way from the manipulative semantics of the OE V+NP+*that* construction, the complement clause nonetheless refers to a deliberate action on the part of the causee and not to a state. This is a relatively late example, from a letter written by Margaret Paston to her husband in 1461, and may reflect a later extension of the construction to non-stative contexts.

4.3.6 *Possible external factors encouraging the development of V+NP+that*
While the essentially language-internal hypothesis which sees ME causative V+NP+ *that* develop from the simple, V+*that* construction (with very similar semantic properties, as we saw in 4.3.1) seems likely, the possible influence of similar structures in other languages in contact with ME, notably French and Latin, should also be taken into account.

To check for the possible influence of Anglo-Norman French, I ran a proximity search for examples where *faire* (all forms given in the *Anglo-Norman Dictionary* (*AND*)) occurred within a 5-word vicinity of *que* in the *Anglo-Norman Dictionary* and in the texts available at the *Ango-Norman On-Line Hub*. No instances of V+NP+*que* came to light, suggesting that the construction, unlike simple V+*que*, did not exist in Anglo-Norman French, at least, and that Anglo-Norman is unlikely to have significantly influenced the development of V+NP+*that* in English.

Latin, on the other hand, possessed a "proleptic accusative" construction, which could be used with causatives, and which bears a number of resemblances to V+NP+*that*, as can be seen from the following examples:[16]

(29) *Fac   me   ut   sciam.*
      make me that I.know
      'Let me know.'   (Marcus Aurelius, *Fronto*, quoted by Halla-Oho 2012: 7)

(30) *Haec me   ut   confidam   faciunt.*
      these me that I.am.confident   make
      'These things make me confident.'
              (Cicero, *Epistulae ad Quintum fratrem*, quoted by Halla-Oho 2012: 6)

(31) *…te faciam   ut   scias*
      you I.will.make that you.know
      'I will let you know'   (Plautus, *Asinaria*, quoted by Bortolussi 1998: 204)

In each of these instances, an NP coreferential with the subject of the finite verb in the complement clause appears as the direct object the matrix causative.

Any possible Latin influence would presumably have been most strongly felt, at least initially, by members of a relatively small educated class, literate to a greater or lesser extent in Latin. Some, at least, would certainly have attained some degree of bilingualism: John Gower, even at a time when English was beginning to reassert itself in literary and official contexts, was literate enough in both Latin and Anglo-Norman French as well as in English to have written major poems in all three languages. Such speakers may have been encouraged to use the form by the fact that a superficially similar structure occurred in Latin. This would not be a case of one language borrowing a structure from another, but rather one in which the Latin syntactic model reinforces the tendency for a ME innovation to become established.

---

16. I would like to thank Elly van Gelderen (p.c.) for pointing out to me the existence of apparently similar constructions not only in Latin but also in ancient Greek.

4.3.7 *Summary*

Given that none of the explanations suggested here seems to be able to account straightforwardly for all of the V+NP+*that* examples I have found, one needs to consider the possibility that several factors may be at work. It would seem that a new V+NP+*that* construction emerged in the late OE/early ME period, at a time when the causative group was undergoing a number of changes. This development was probably influenced most directly by the existing V+*that* and V+NP+Pred constructions, with which it shared a number of semantic properties, but may also have been encouraged by other, superficially similar constructions, such as Fischer's topicalized structures discussed in Section 4.3.4, and even Warner's CLAN sentences found with other verbs.[17] In terms of the "network" approach proposed by Traugott & Trousdale (2013), this would seem to be a case of constructionalization, a new node being created within a constructional network, linked semantically and formally (at least in part) to V+*that* and V+NP+Pred, and formally to the topicalization and, superficially at least, to the CLAN construction.

5. Conclusion

The facts presented in this paper allow us to sketch the probable history of finite causative complements in ME. The two types of finite complement appear to have had different origins. The simple V+*that* structures found in early ME were probably survivors from OE, occurring in much the same kinds of second causative

---

17. Hendrik de Smet (p.c.) suggests that causative V+NP+*that* may initially have arisen as a form of constructional blend, between competing V+*that* and the *to*-infinitive VOSI[to] constructions. In other words, two distinct linear sequences, V+*that*+NP+finite verb and V+NP+ *to*-infinitive would have blended to produce V+NP+*that*+finite verb. It is by no means inconceivable that VOSI[to] could have influenced the development of V+NP+*that*, especially with the verb *make*. Both constructions were associated with second causative semantics, and would presumably have found themselves in competition, at least to an extent.

According to Visser (1973), some form of blending, at least, did take place in causative complement clauses, as in the following example:

> *Allas! the shorte throte, the tendre mouth, /* **Maketh that** *est and west and north and south, / In erthe, in eir, in water,* **men to swynke** */ To gete a glotoun deyntee mete and drynke!*
> 'Alas! The short throat, the tender mouth, make men labor in the east, the west, the north and the south, in earth, in air, and in water, to get dainty food and drink for a glutton!'
> (Chaucer, *Pardoner's Tale*, 517–520)

Visser quotes several examples, with a variety of verbs, considering the construction to be a blend: "This anacoluthic construction is clearly a hybrid of a VOSI and a *that* clause" (1973: § 2060). For discussion of syntactic blending as a mechanism of innovation, see De Smet (2014: 23–48).

environment as OE *(ge)don* with a *that*-complement. Rather than dying out completely, as infinitive complements replaced finite ones in most contexts, finite causative complements remained at least marginally productive throughout the ME period, to the extent that they could even be associated with new causative verbs, such as *cause*, in much later ME. The resistance of V+*that* may well be due, at least in part, to contact with Anglo-Norman, in which causative *faire* was used with a finite *que*-complement. A new form, V+NP+*that*, appeared in late OE or early ME: this construction had little to do with the superficially similar form frequent in OE with manipulative verbs. The semantics of the new variant were essentially of the same second causative type as those of the simple V+*that* construction. The new form too remains productive, albeit marginally, until the end of the ME period and even beyond, examples being found until EModE. The emergence of the new construction is probably a language-internal development, primed by a number of existing constructions to which it was syntactically or semantically related.

Some of the questions raised in this paper are still to be answered. For example, if we accept the hypothesis that language contact played a role in keeping the finite constructions alive, the question of exactly how ME was influenced, both by Anglo-Norman and by Latin needs to be further explored. In addition, the relatively marginal status of both V+*that* and V+NP+*that* is intriguing. Both constructions remain somewhat peripheral throughout the ME period, and yet sufficently productive for *cause*, a comparatively late addition to the group of causative verbs to begin to appear in them in later ME. Why they should have conserved their essentially marginal status, never becoming fully established syntactic options nor falling into obsolecence, remains to be determined.

Acknowledgements

References

Primary sources

ÆCHom 2
Thorpe, Benjamin (ed.). 1844–46. *The Sermones Catholici or Homilies of Ælfric*, Vol. II. London: Ælfric Society.

Alphabet of Tales
Banks, Mary (ed.). 1904–05. *Alphabet of Tales: An English 15th Century Translation of the Alphabetum Narrationum of Etienne de Besançon* [Early English Text Society Original Series 126–127]. London: Kegan Paul, Trench, Trübner & Co.

Ancrene Wisse
Hasenfratz, Robert (ed.). 2000. *Ancrene Wisse* [TEAMS Middle English Texts Series]. Kalamazoo MI: Medieval Institute Publications. Online edn. <http://d.lib.rochester.edu/teams/publication/hasenfratz-ancrene-wisse> (27 June 2017).

Athelston
Herzman, Ronald, Drake, Graham & Salisbury, Eva (eds). 1997. *Four Romances of England* [TEAMS Middle English Texts Series]. Kalamazoo MI: Medieval Institute Publications. Online edn. <http://d.lib.rochester.edu/teams/text/salisbury-athelston> (27 June 2017).

Brut
Brook, George & Leslie, Roy F. (eds). 1963, 1978. *Laȝamon: Brut*. (Ms. Cotton Caligula A.IX) [Early English Text Society Original Series 250, 277]. Oxford: OUP. Online edn, University of Michigan Humanities Text Initiative. <http://name.umdl.umich.edu/LayCal> (27 June 2017).

Canterbury Tales
Benson, Larry (ed.). 1988. *The Riverside Chaucer*. Oxford: OUP.

Chaucer, prose texts
Benson, Larry (ed.). 1988. *The Riverside Chaucer*. Oxford: OUP.

Cleanness, Patience, Pearl
Morris, Richard (ed.). 1869. *Early English Alliterative Poems of the Fourteenth Century* [Early English Text Society Original Series 1]. London: Trübner & Co.

Confessio Amantis
Macaulay, George (ed.). 1899–1902. *The Complete Works of John Gower*. Oxford: Clarendon Press.

Dame Siriþ
Salisbury, Eva (ed.). 2002. *The Trials and Joys of Marriage* [TEAMS Middle English Texts Series]. Kalamazoo MI: Medieval Institute Publications. Online edn. <http://d.lib.rochester.edu/teams/text/salisbury-trials-and-joys-dame-sirth> (27 June 2017).

Hali Meiðhad

Millet, Bella (ed.). 1972. *Hali Meiðhad* [Early English Text Society Original Series 284]. Oxford: OUP.

Havelok

Skeat, Walter (ed.). 1868. *The Lay of Havelok the Dane* [Early English Text Society Extra Series 4]. London: N. Trübner & Co.

Heptateuch

Marsden, Richard (ed.). 2008. *The Old English Heptateuch and Ælfric's Libellus de veteri testamento et novo* [Early English Text Society 330]. Oxford: OUP.

Holy Rood Tree

Napier, Arthur (ed.). 1894. *The History of the Holy Rood-tree: A Twelfth Century Version of the Cross Legend* [Early English Text Society Original Series 103]. London: K. Paul, Trench, Trübner & Co.

King Horn

Hall, Joseph (ed.). 1901. *King Horn; A Middle-English Romance*, edited from the manuscripts by Joseph Hall (MS Gg IV 27.2). Oxford: Clarendon Press. Online edn, University of Michigan Humanities Text Initiative. <http://name.umdl.umich.edu/ACN1637.0001.001> (27 June 2017).

Lambeth Homilies

Morris, Richard (ed.). 1867–68. *Old English Homilies and Homiletic Treatises of the Twelfth Century* [Early English Text Society 29, 34]. London: Trubner & Co. Online edn, University of Michigan Humanities Text Initiative. <http://name.umdl.umich.edu/AHA2688.0001.001> (27 June 2017).

Mandeville's Travels

Hamelius, Paul (ed.). 1919, 1923. *Mandeville's Travels*, translated from the French of Jean d'Outremeuse [Early English Text Society Original Series 153, 154]. London: K. Paul, Trench, Trübner & Co. Online edn, University of Michigan Humanities Text Initiative. <http://name.umdl.umich.edu/aeh6691> (27 June 2017).

Morte d'Arthur

Vinaver, Eugène (ed.). 1971. *Malory: Works*, 2nd edn. Oxford: OUP.

Sommer, Heinrich Oskar (ed.). 1899. *Le Morte Darthur by Syr Thomas Malory; The Original Edition of William Caxton Now Reprinted and Edited with an Introduction and Glossary by H. Oskar Sommer; With an Essay on Malory's Prose Style by Andrew Lang*. London: David Nutt. Online edn, University of Michigan Humanities Text Initiative. <http://name.umdl.umich.edu/MaloryWks2> (27 June 2017).

Octovian

McSparran, Frances (ed.). 1986. *Octovian, Edited from Lincoln, Dean and Chapter Library, MS 91 and Cambridge, University Library, MS Ff.2.38* [Early English Text Society Original Series 289]. Online edn, University of Michigan Humanities Text Initiative. <http://name.umdl.umich.edu/OctC; http://name.umdl.umich.edu/OctL> (27 June 2017).

Owl & Nightingale

Stanley, Eric (ed.). 1960. *The Owl and the Nightingale*. London: Nelson.

Paston Letters

Davis, Norman (ed.). 1971. *The Paston Letters and Papers of the Fifteenth Century*. Oxford: Clarendon Press.

Pepys, Diary

Latham, Robert & Matthews, William (eds). 1970–83. *The Diary of Samuel Pepys*, 11 Vols. London: G. Bell.

Peterborough Chronicle

Jebson, Tony (ed.). 1996. *The XML Edition of the Anglo-Saxon Chronicle* (MS 'E'). Online edn. <http://asc.jebbo.co.uk/e/e-L.html> (27 June 2017).

Piers Plowman

Schmidt, Aubrey (ed.). 1978. *The Vision of Piers Plowman: A Critical Edition of the B-text*. London: J.M. Dent and E.P. Dutton.

Prose Rule of St. Benet

Kock, Ernst (ed.). 1902. *Three Middle English Versions of the Rule of St. Benet and two Contemporary Rituals for the Ordination of Nuns* [Early English Text Society Original Series 120]. London: K. Paul, Trench, Trübner & Co.

Richard Rolle, Prose and Verse

Ogilvie-Thomson, Sarah (ed.). 1988. *Richard Rolle: Prose and Verse* [Early English Text Society 293]. Oxford: OUP.

Sawles Warde

Morris, Richard (ed.). 1867–68. *Old English Homilies and Homiletic Treatises of the Twelfth Century* [Early English Text Society 29, 34]. London: Trubner & Co. Online edn, University of Michigan Humanities Text Initiative. <http://name.umdl.umich.edu/AHA2688.0001.001> (27 June 2017).

Siege of Jerusalem

Kolbing, Eugen & Day, Mabel (eds). 1932. *The Siege of Jerusalem, Edited from MS Laud. Misc. 656* [Early English Text Society 188]. London: OUP.

Sir Gawain & the Green Knight

Morris, Richard (ed.). 1869. *Sir Gawayne and the Green Knight: An Alliterative Romance Poem* [Early English Text Society Original Series 4]. London: Trübner & Co. Online edn, Project Gutenberg. <http://www.gutenberg.org/files/14568/14568-h/14568-h.htm. (27 June 2017).

Sir Orfeo
Bliss, Alan (ed.). 1966. *Sir Orfeo*, 2nd edn. Oxford: Clarendon Press.

Troilus & Criseyde
Benson, Larry (ed.). 1988. *The Riverside Chaucer*. Oxford: OUP.

Vices & Virtues
Holthausen, Ferdinand (ed.). 1888–1921. *Vices and Virtues, Being a Soul's Confessions of its Sins* [Early English Text Society 89, 159]. London: OUP.

York Mystery Plays
Beadle, Richard (ed.). 1982. *The York Plays*. London: E. Arnold. Online edn, University of Michigan Humanities Text Initiative. <http://name.umdl.umich.edu/York> (27 June 2017).

Dictionaries and electronic corpora

AND
Anglo-Norman Dictionary, available at *the Anglo-Norman On-Line Hub* <http://www.anglo-norman.net>

OED
Simpson, John & Weiner, Edmund (eds). 1989. *The Oxford English Dictionary*, 2nd edn, 20 Vols. Oxford: Clarendon Press.

MED
Kurath, Hans (ed.). 1954. *Middle English Dictionary*. Ann Arbor MI: University of Michigan Press. Online edn. <http://quod.lib.umich.edu/m/med/>

ANHT
Anglo-Norman Hub Textbase. <http://www.anglo-norman.net>

HC
Helsinki Corpus TEI XML Edition. 2011. 1st edn. Designed by Alpo Honkapohja, Samuli Kaislaniemi, Henri Kauhanen, Matti Kilpiö, Ville Marttila, Terttu Nevalainen, Arja Nurmi, Matti Rissanen and Jukka Tyrkkö. Implemented by Henri Kauhanen and Ville Marttila. Based on *The Helsinki Corpus of English Texts* (1991). Helsinki: The Research Unit for Variation, Contacts and Change in English (VARIENG), University of Helsinki.

Secondary sources

Bertolussi, Bernard. 1998. *Facite ventvm vt gavdeam*: Quelques phénomènes d'ambiguïté syntaxique. In *Estudios de linguistica latina*, Benjamin García-Hernández (ed.), 203–216. Madrid: Ediciones Clásicas.

Bock, Hellmut. 1931. Studien zum präpositionalen Infinitiv und Akkusativ mit dem *to*- Infinitiv. *Anglia* 55: 114–249.

Denison, David. 1993. *English Historical Syntax: Verbal Constructions*. London: Longman.

De Smet, Hendrik. 2014. Does innovation need reanalysis? In *Usage-based Approaches to Language Change* [Studies in Functional and Structural Linguistics 69], Evie Coussé & Ferdinand von Mengden (eds), 23–48. Amsterdam: John Benjamins.

Fischer, Olga. 1987. Some remarks on the analysis of perception verb complements in Middle English: A reply. *Kwartalnik Neofilologiczny* XXXIV: 57–67.

Fischer, Olga. 1990. *Syntactic Change and Causation: Developments in Infinitival Constructions in English* [Amsterdam Studies in Generative Grammar 2]. Amsterdam: University of Amsterdam.

Halla-aho, Hilla. 2012. A historical perspective on Latin proleptic accusatives. *De Lingua Latina* 7. Online edn. <http://www.paris-sorbonne.fr/IMG/pdf/Halla-aho_Proleptic_Accusatives_relecture-final.pdf> (27 June 2017).

Ikegami, Yoshihiko. 1981. On the development of the causative use of the verbs *do* and *make* in English. In *Eigo no rekishi to kozo: Festschrift for Professor Miyabe*, Yoshio Terasawa et al. (eds), 85–99. Tokyo: Kenyusha Shuppani.

Ingham, Richard. 2012. Middle English and Anglo-Norman in contact. *Bulletin des Anglicistes Médiévistes* 81: 1–24.

Karttunen, Lauri. 1971. Implicative verbs. *Language* 47: 340–358. doi:10.2307/412084

Kopytko, Roman. 1985. Some observations on the possible interrelationship between synchronic and diachronic data in syntactic analysis. *Kwartalnik Neofilologiczny* XXXII: 27–32.

Kroch, Anthony & Taylor, Ann. 1997. Verb movement in Old and Middle English. In *Parameters of Morphosyntactic Change*, Ans van Kemenade & Nigel Vincent (eds), 297–325. Cambridge: CUP.

Los, Bettelou. 2005. *The Rise of the* To-*Infinitive*. Oxford: OUP.
  doi:10.1093/acprof:oso/9780199274765.001.0001

Lowrey, Brian. 2002. Les verbes causatifs en anglais: Une étude diachronique du moyen-anglais à l'anglais moderne. PhD dissertation, Université de Lille III.

Lowrey, Brian. 2010. Causative verbs in West Saxon Old English. *Bulletin des Anglicistes Médiévistes* 78: 57–88.

Lowrey, Brian. 2012. Early English causative constructions and the second agent factor. In *Outposts of Historical Corpus Linguistics: From the Helsinki Corpus to a Proliferation of Resources* [Studies in Variation, Contacts and Change in English 10], Jukka Tyrkkö, Matti Kilpiö, Terttu Nevalainen & Matti Rissanen (eds.). Helsinki: Research Unit for Variation, Contacts, and Change in English. Online edn: <http://www.helsinki.fi/varieng/journal/volumes/10/> (27 June 2017).

Lowrey, Brian. 2013a. Syntactic and semantic variation: Middle English causatives and complement types. In *Periphrasis, Replacement and Renewal: Studies in English Historical Linguistics*, Irén Hegedűs & Dora Pődör (eds), 92–123. Newcastle upon Tyne: Cambridge Scholars.

Lowrey, Brian. 2013b. The Old English causative verb *hatan* and its demise. *Token* 2: 23–43.

Manabe, Kazumi. 1989. *The Syntactic and Stylistic Development of the Infinitive in Middle English*. Fukuoaka: Kyushu University Press.

Mustanoja, Tauno. 1960. *A Middle English Syntax, Part one: Parts of speech*. Helsinki: Société Néophilologique.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan. 1985. *A Grammar of Contemporary English*. London: Longman.

Royster, James F. 1918. The causative use of the verb *hatan*. *Journal of English and Germanic Philology* 17: 82–93.

Royster, James F. 1922. Old English causative verbs. *Studies in Philology* 19: 328–356.

Short, Ian. 2007. *Manual of Anglo-Norman* [Occasional Publications Series 8]. London: Anglo-Norman Text Society, Birkbeck College, University of London.

Terasawa, Jun. 1985. The historical development of the causative use of the verb *make* with an infinitive. *Studia Neophilologica* 57: 133–143. doi:10.1080/00393278508587913

Traugott, Elisabeth C. & Trousdale, Graeme. 2013. *Constructionalization and Constructional Changes*. Oxford: OUP. doi:10.1093/acprof:oso/9780199679898.001.0001

Vendler, Zeno. 1967. Verbs and times. In *Linguistics in Philosophy*, Zeno Vendler (ed.), 97–121. Ithaca CA: Cornell University Press.

Visser, Frederikus Theodorus. 1973. *An Historical Syntax of the English Language*. Leiden: E. J. Brill.

Warner, Anthony. 1982. *Complementation in Middle English and the Methodology of Historical Syntax*. London: Croom Helm.

Appendix 1

Occurrences of finite clausal causative complements, individual texts.

Table 3. Finite and non-finite causative complements, *Vices & Virtues*

Vices & Virtues	*do*	*make*	*gar*	*cause*	*let*
Finite complement clauses	**3**	**1**			
+*that*	1	0			
+NP+*that*	2	1			
Infinitive complement clauses	**14**	**1**			
+S	12	1			
−S	2	0			

Table 4. Finite and non-finite causative complements, *Lambeth Homilies*

Lambeth Homilies	do	make	gar	cause	let
Finite complement clauses	**6**	**1**			
+*that*	5	1			
+NP+*that*	1	0			
Infinitive complement clauses	**1**	**6**			
+S	1	6			
−S	0	0			

Table 5. Finite and non-finite causative complements, *Havelok*

Havelok	do	make	gar	cause	let
Finite complement clauses	**2**	**0**	**0**		
+*that*	1	0	0		
+NP+*that*	1	0	0		
Infinitive complement clauses	**66**	**24**	**5**		**2**
+S	21	19	4		
−S	45	5	1		2

Table 6. Finite and non-finite causative complements, *Owl and Nightingale*

Owl & Nightingale	do	make	gar	cause	let
Finite complement clauses	**1**	**0**			
+*that*	1	0			
+NP+*that*	0	0			
Infinitive complement clauses	**2**	**3**			**2**
+S	2	2			
−S	0	1			2

Table 7. Finite and non-finite causative complements, *Piers Plowman*

Piers Plowman	do	make	gar	cause	let
Finite complement clauses	**0**	**2**	**0**		
+*that*	0	1	0		
+NP+*that*	0	1	0		

Table 7. (*continued*)

Piers Plowman

	do	make	gar	cause	let
Infinitive complement clauses	**46**	**38**	**10**		
+S	35	38	9		
−S	11	0	1		

Table 8. Finite and non-finite causative complements, *Confessio Amantis*

Confessio Amantis

	do	make	gar	cause	let
Finite complement clauses	**2**	**7**		**1**	
+*that*	2	5		1	
+NP+*that*	0	2		0	
Infinitive complement clauses	**45**	**67**		**3**	**37**
+S	31	65		1	
−S	14	2		2	37

Table 9. Finite and non-finite causative complements, *Chaucer, Prose Texts*

Chaucer: prose texts

	do	make	gar	cause	let
Finite complement clauses	**0**	**7**			
+*that*	0	7			
+NP+*that*	0	0			
Infinitive complement clauses	**3**	**31**			**5**
+S	3	30			
−S	0	1			5

Table 10. Finite and non-finite causative complements, *Troilus and Criseyde*

Troilus & Criseyde

	do	make	gar	cause	let
Finite complement clauses	**1**	**1**		**0**	
+*that*	1	0		0	
+NP+*that*	0	1		0	
Infinitive complement clauses	**29**	**14**		**5**	**1**
+S	29	14		5	
−S	0	0		0	1

Table 11. Finite and non-finite causative complements, *Canterbury Tales*

Canterbury Tales	do	make	gar	cause	let
Finite complement clauses	**1**	7		0	
+*that*	1	3		0	
+NP+*that*	0	4		0	
Infinitive complement clauses	**39**	**89**		**6**	**20**
+S	19	86		4	
−S	20	3		2	20

Table 12. Finite and non-finite causative complements, *Morte d'Arthur*

Morte d'Arthur	do	make	gar	cause	let
Finite complement clauses	**0**	**0**	**0**	3	
+*that*	0	0	0	0	
+NP+*that*	0	0	0	3	
Infinitive complement clauses	**24**	**168**	**7**	**37**	**149**
+S	10	156	1	37	
−S	14	12	6	0	149

Table 13. Finite and non-finite causative complements, *Paston Letters*

Paston Letters	do	make	gar	cause	let
Finite complement clauses	**0**	**0**		2	
+*that*	0	0		1	
+NP+*that*	0	0		1	
Infinitive complement clauses	**98**	**19**		**100**	
+S	29	19		99	
−S	69	0		1	

Appendix 2

Occurrences of V+*that* and V+NP+*that* in ME texts

V+ *that*

and swa ðu miht **don** ðat all ðat euel ðe mai wanden te michele gode
Vices & Virtues 29: 13

forðon þe he **dude** þet heo weren birnende on godes willan. and bodiende umbe godes riche   *Lambeth Homilies* 95: 12

On is icweðen. Gula. þet is ȝifernesse on englisc. þeo **deð** þet mon et er timan. and drinceð. oðer eft to muchel nimeð on ete oðer on wete   *Lambeth Homilies* 103: 2

Ira. þet is on englisc wemodnesse. heo **deð** þet þe mon ne ah his modes iwald
Lambeth Homilies 103: 8

Ðurh þisse tacne **makede** Moyses þet ðet weter of egipte wes liðe and swete þan folce of israel.   *Lambeth Homilies* 129: 2

Ich schal mid one bare worde / **Do** þat þi speche w[u]rþ forwurþe
Owl & Nightingale 547–548

Meschief it **maketh** thei be so meke nouthe   *Piers Plowman* VI: 205

Among the men lacke of manhode / In Mariage upon wifhode / **Makth** that a man himself deceiveth   *Confessio Amantis* V: 455- 457

Bot Jelousie of his untrist / **Makth** that full many an harm arist
Confessio Amantis V: 717–718

(He mot) **Do** that ther be fixacion / With tempred hetes of the fyr
Confessio Amantis IV: 2520–2521

For thanne upon Astronomie / Of due constellacion / Thou makst prolificacion, / And **dost** that children ben begete   *Confessio Amantis* IV: 3246–3249

That crualte hath no good ende; / Bot Pite, hou so that it wende, / **Makth** that the god is merciable, / If ther be cause resonable   *Confessio Amantis* VII: 3518–3521

And thus the foles evidence, / Which was of goddes grace enspired, / **Makth** that good conseil was desired   *Confessio Amantis* VII.4002–4004

His eloquence and his facounde / Of goodly wordes whiche he tolde, / Hath **mad** that Anthenor him solde / The toun, which he with tresoun wan
Confessio Amantis VII: 1560–1563

The lond is ful of maintenue, / Which **causith** that the comune right / In fewe contrees stant upright   *Confessio Amantis* VIII: 3012–3014

Of Accidie comth first that a man is anoyed and encombred for to doon any goodnesse, and **maketh** that God hath abhomynacion of swich Accidie, as seith Seint John
Chaucer Prose, *Parson's Tale*, 311: 686

Usage of labour is a greet thyng, for it **maketh**, as seith Seint Bernard, the laborer to have stronge armes and harde synwes; and slouthe maketh hem feble and tendre
Chaucer Prose, *Parson's Tale*, 311: 689

For it bireveth hym the love that men to hym owen, and turneth it bakward agayns alle resoun, and **maketh** that the avaricious man hath moore hope in his catel than in Jhesu Crist, and dooth moore observance in kepynge of his tresor than he dooth to the service of Jhesu Crist   Chaucer Prose, *Parson's Tale*, 313: 745–746

Thanne mai nat richesses **maken** that a man nys nedy, ne that he be suffisaunt to hym-self   Chaucer Prose, *Boece*, 424: 52–53

[he] hadde **maked** that the hare was nat agast of the hound, whiche was plesed by his song   Chaucer Prose, *Boece*, 439: 11–12

In loves termes; hold of thi matere / The forme alwey / And **do** that it be lik
                    Troilus & Criseyde II: 1037–40

This **maketh** that ther ben no fayeryes   *Canterbury Tales* 117: 872

The clennesse and the fastynge of us freres / **Maketh** that crist accepteth oure preyeres
                    Canterbury Tales 131: 1883–1884

He **made** that the ryver was so smal / That wommen myghte wade it over al
                    Canterbury Tales 133: 2083–2084

Thanne help me, lord, tomorwe in my bataille, / For thilke fyr that whilom brente thee, / As wel as thilke fyr now brenneth me, / And **do** that I tomorwe have victorie
                    Canterbury Tales 57: 2402–2405

But I conceyve now the trought, and þat sikenesse **caused** that she [be] absent here
                    Paston Letters 369: 8–9

V+NP+*that*

Ðies **dieð** ðe manne ðe ðurh hire is beswiken ðat he twifealdeð his senne
                    Vices & Virtues 5: 20

ðie[s] iwerȝede gast me hafð **idon**, ðat ic am swiðe forȝelt aȝeanes mine laferde god almihtin.   *Vices & Virtues* 7: 6

Hie nis naht to laten, for ðan hie **makeð** ðanne man ðe godd his to-sant þat he þur-wuneð on his godnesse   *Vices & Virtues* 149: 28

forðon þet heo **deð** þere monnan heortan þet he onlihteð mid his ȝife. þet heo beoð liðe þurh un-cladnesse*. [unlaðnesse.]   *Lambeth Homilies* 97: 15

Hwan he hauede **don** him for drede, / Þat he sholde hire spusen, and fede, / And þat she sholde til him holde, / Þer weren penies þicke tolde   *Havelok* 1169–1172

Mighte no lesynges, ne lyere, ne los of worldly catel / **Maken** hym, for any mournynge, that he nas murie in soule   *Piers Plowman* XIX: 294–295

Anon the wylde loves rage, / In which noman him can governe, / Hath **mad** him that he can noght werne   *Confessio Amantis* I: 2620–2622

Bot thilke unkynde Jelousie, / Which everemor the herte opposeth, / **Makth** Vulcanus that he supposeth / That it is noght wel overal   *Confessio Amantis* V: 670–673

The drede of lesyng **maketh** hym that he / May in no perfyt selynesse be
                    Troilus & Criseyde III-827–31

Myn heed is toty of my swynk to-nyght, / That **makes** me that I ga nat aright
                    Canterbury Tales 83: 4253–4254

And yet in bacon hadde I nevere delit; / That **made** me that evere I wolde hem chide.

Canterbury Tales 110: 418–419

I folwed ay myn inclinacioun / By vertu of my constellacioun; / That **made** me I koude noght withdrawe / My chambre of venus from a good felawe

Canterbury Tales 113: 615–618

For which, where as his peple therbifore / Hadde loved hym wel, the sclaundre of his diffame / **Made** hem that they hym hatede therfore   *Canterbury Tales* 146: 729–731

this rynge is passynge mete for me for it wille torne al manere of lykenes that I am in and that shalle **cause** me that I shall not be knowen   *Morte d'Arthur* 257: 17

Alle that knewe I afore hand sayd kynge Arthur and that **caused** me I wold not suffer you to haue adoo atte grete Iustes for I aspyed said kynge Arthur whan he cam in tyl his lodgynge   *Morte d'Arthur* 750: 13

And for your presumpcyon to take vpon you in dedely synne for to be in his presence where his flesshe and his blood was that **caused** you ye myghte not see hit with worldly eyen for he wille not appiere where suche synners ben   *Morte d'Arthur* 640: 2

they bothe told me þat ye schold veryly a ben at home be-for Crystmas, and that **causyd** me that I wrot not to yow non answer   *Paston Letters* 276: 11–12

Causative *make* and its infinitival complements in Early Modern English

Yoko Iyeiri
Kyoto University

The present paper discusses the variability between *to-* and bare infinitives in the complement of causative *make* in Early Modern English, with a focus on *make* used in the active voice. Although bare infinitives are almost exclusively used with active causative *make* in Present-day English, this has not always been the case in the history of English. Throughout the Early Modern English period, *to*-infinitives were still encountered to a noticeable extent. Moreover, various linguistic factors were relevant to the choice of *to-* and bare infinitives. The present paper examines complexity in general, the objects/causees of *make*, coordination of infinitives, and lexical items used as infinitives.

Keywords: causative *make*, complementation, bare infinitives, *to*-infinitives, complexity

1. Introduction

In Present-day English, causative *make* is usually followed by bare infinitives when it is used in the active voice and *to*-infinitives when used in the passive voice. See examples (1) and (2), both from the *British National Corpus* (*BNC*):[1]

(1) In four areas local conditions *made* him *favour* a unitary approach.
(*BNC*, *Local Government and Urban Politics* by William Hampton)

(2) …and that was unfortunately how I *was made to feel* for much of the time.
(*BNC*, *Part of the Furniture* by Michael Falk)

In the following example, however, active *make* is followed by a *to*-infinitive:

(3) What inversion of values *makes* us *to look* upon such aberrations as though they were a reflection of natural laws? (*BNC*, *New Statesman and Society*)

---

1. Examples (1), (2), and (3) have been cited from: <http://corpus.byu.edu/bnc/> (3 May 2016).

© 2018 John Benjamins Publishing Company

The use of *to*-infinitives in the active voice, as illustrated by (3), is exceptional in Present-day English, and often restricted to specific circumstances. *The Oxford English Dictionary* (s.v. *make*) maintains that the use of *to*-infinitives with active *make* is somewhat archaic. Andersson (1985: 149) argues that *to*-infinitives tend to occur in "religiously coloured" language.

In the history of English, however, the distinction between "active *make* + bare infinitive" and "passive *make* + *to*-infinitive" has not always been as clear as it is today. The key period in terms of the establishment of the bare infinitive with active *make* seems to be from later Middle English to Early Modern English. There are a number of studies dealing with the occurrence of *to*- and bare infinitives with causative *make* in the 14th and 15th centuries, which all indicate that *to*-infinitives were still commonly used with active *make* in late Middle English (see Sugiyama 1988; Sawada 1997; Iyeiri 2012; and Lowrey 2013, among many others).[2] The use of *to*-infinitives, however, decreases in the active voice in the course of the Early Modern English period. According to Fanego (1994: 196–197), the dominant position of bare infinitives was largely established by the time of Shakespeare and Dryden.

In the following sections, I will account for the sharp decline of *to*-infinitives with causative *make* in Early Modern English. After detailing the data and methods of the present study in Section 2, I will deal with the overall expansion of bare infinitives in the Early Modern English period in Section 3. I will then explore various factors which are possibly relevant to the choice of infinitival forms in Section 4, and conclude the entire discussion in Section 5.

2.   Data and methods

This study is based on *EMEPS* (*Early Modern English Prose Selections*, version 1), which I have compiled for various research purposes, and not necessarily for the present paper alone. It consists of Early Modern English prose texts selected from the *Early English Books Online*, which I have divided into A-texts (about four million words in total) and B-texts (also about four million words), as shown in Table 1:[3]

---

2.   Strictly speaking, Middle English yields *for to*-infinitives as the third possibility. See the following example, which is quoted from Iyeiri (2012: 59) and ultimately from *ICAMET* (*Innsbruck Computer Archive of Machine-Readable English Texts*): "He ouyrcam þe kyng Lettow, and mad him *for to fle*" (*ICAMET*, *John Capgrave*). The Early Modern English texts investigated in the present study no longer yield any examples of *for to*-infinitives.

3.   For the *Early English Books Online*, see: <http://eebo.chadwyck.com/home> (3 May 2016).

Table 1. *EMEPS* (A-texts, B-texts)

	Files	Dates	Number of texts	Approximate size (words)
A-texts	A16th–1	1501–1550	43	1,007,800
	A16th–2	1551–1600	53	1,002,400
	A17th–1	1601–1650	95	1,002,200
	A17th–2	1651–1700	75	1,001,200
B-texts	B16th–1	1501–1550	39	1,009,600
	B16th–2	1551–1600	34	1,003,300
	B17th–1	1601–1650	39	1,004,600
	B17th–2	1651–1700	43	1,007,400

Since causative *make* is a fairly common lexical item, I will use A-texts only in the present study, with the assumption that four million words will yield a sufficient number of examples.[4]

   As expected, causative *make* is followed by infinitives with or without *to* in *EMEPS*, as illustrated by (4) and (5):

(4)   It *made* her *to stagger*, …   (William Camden, 1624)[5]

(5)   I am confident, all the world could not *make* him *commit* a known sinne.
                                                                    (George Sondes Feversham, 1655)

The competition is virtually confined to these two types of complements, although examples of *that*-clauses as complements of *make* are also attested in *EMEPS*, as in:

(6)   And thys co~iecture *maketh* me, *that* I thynke the yeares fro the celestiall Adam
       vnto the fyerve floude in the seconde coniecture, and the iubile yeares of the
       church in the thirde coniecture must not be rekened fro the byrthe of our
       Lorde but from hys resurreccyon.   (Andreas Osiander, 1548)

That-clauses as complements of causative *make* are, however, already marginal and close to extinction in the Early Modern English period, occurring only eight times

---

4.   I have divided the corpus into A-texts and B-texts, assuming that some research topics will require eight million words while others may need only four million words. Since the texts in A and B have been selected essentially on the same principle, the B-texts alone could equally well have been used for this study. The gist here is that I have decided to use four million words instead of eight million words (A and B combined) for the purpose of the present study. For further details and the titles of the texts included in *EMEPS*, see Iyeiri (2011).

5.   Unless otherwise stated, the citations in the present paper are from *EMEPS*, and therefore ultimately from the *Early Modern English Books Online*. Since titles of books published in the relevant period tend to be long, often too long to quote, this paper indicates the author and the date of publication only. The emphasis in the quotations is mine.

in the data explored in the present study.[6] The discussion below is, therefore, restricted to causative *make* with *to-* and bare infinitives.

Furthermore, I will restrict this study to the active voice of causative *make*, for two reasons. Firstly, relevant examples in the passive are infrequent: the texts under analysis provide only 81 passive voice examples as against 1,198 examples in the active voice.[7] Secondly and more importantly, causative *make* in the passive is essentially different from causative *make* in the active voice, in that the former occurs with *to*-infinitives even in contemporary English. It is relevant to state in passing, though, that passive *make* also shows variation in complementation patterns in Early Modern English. See examples (7) and (8), where passive *make* is followed by infinitives with and without *to*:

(7)   For we are *made* in Chryste Iesu *to do* the good workes whyche God hath prepared that we shoulde walke in them.   (Stephen Gardiner, 1546)

(8)   …and how can that be *made appear*, but by being governed by his Laws?
(James Davies, 1679)

Although about 80% of the relevant examples of passive *make* in *EMEPS* take *to*-infinitives, those with bare infinitives are also attested, as illustrated by (8). Thus it appears that the historical path followed by causative *make* is a path along which the active-voice and passive-voice uses have gradually diverged: while both infinitival forms were still attested with active and passive *make* in the Early Modern English period, bare infinitives have become increasingly common with active *make* and *to*-infinitives with passive *make* in the course of the history of English.

To return to the research methodology of the present study, when two or more infinitival complements are coordinated, the first infinitive alone will be counted, except in Section 4.4, which specifically treats coordinated infinitives. In example (9), for instance, *conclude* will be counted but not *assent*, although both are complements of *make*:

(9)   now urges nature so Artificially, that he *makes* her *conclude* & *assent* to work wonders
(John Gaule, 1646)

---

6.   Seven of the eight examples are found in the 16th century. Rohdenburg (1995: 371) also provides statistics to show how marginal finite clauses are as complements of *make*, *cause*, *bid*, and *help* in Early Modern English. The decrease of *that*-clauses as complements is a general phenomenon observed with a number of English verbs during this period (see also Los 2005 and Iyeiri 2010 among others).

7.   The totals here include only those examples followed by infinitives.

Finally, examples like (10), in which *make* is accompanied by another verb will be excluded from the analysis:

(10) Item, whether they *make* or *cause* to be made their monethly and quarterly Sermons, instruct and examine the youth of theyr Parishes in the Catechisme at Euenyng Prayer on Sondayes and holydayes.   (Matthew Parker,1575)

In this example, *make* is accompanied by *cause*, which may have affected the choice of the infinitival form that follows.[8]

3. Causative *make* plus *to-* and bare infinitives in Early Modern English: A historical overview

As stated above, active causative *make* was still combined with *to-* and bare infinitives in the Early Modern English period, though this variation was significantly reduced in the course of the 17th century and the use of *to*-infinitives with causative *make* was nearly absent in the 18th century (Fanego 1994: 196–197; Rohdenburg 1996: 155–159; Rohdenburg 2000: 28–30; Schlüter 2005: 189–197). Table 2 shows the rise of bare infinitives and the fall of *to*-infinitives in the data concerned:

Table 2. Active causative *make* followed by *to*-infinitives and bare infinitives in *EMEPS*

	To-infinitives	Bare infinitives	Totals
A16th–1 (1501–1550)	151 (68.6%)	69 (31.4%)	220
A16th–2 (1551–1600)	115 (36.5%)	200 (63.5%)	315
A17th–1 (1601–1650)	115 (36.1%)	204 (63.9%)	319
A17th–2 (1651–1700)	60 (17.4%)	284 (82.6%)	344

The general trend is clear and straightforward in *EMEPS*. In the first half of the 16th century, bare infinitives were much less frequent than *to*-infinitives with active causative *make*.[9] It is only from the middle of the 16th century that bare infinitives

---

8. Examples like this are rare in any case in the data under consideration.

9. This tendency may be slightly exaggerated, since A16th–1 includes a few texts written in Middle English but printed in 1501–1550. This is comparable to the Early Middle English period, when a number of Anglo-Saxon works were transcribed by Middle English scribes. Texts of this kind are not excluded from the analysis, partly because the principle of *EMEPS* is to classify texts according to their dates of printing – this applies not only to A16th–1 but also to A16th–2, A17th–1, and A17th–2, and also because it is only natural for texts printed around 1501–1510 to show some features of continuity between Middle English and Early Modern English.

outnumbered *to*-infinitives. The shift from *to*-infinitives to bare infinitives was not yet complete even towards the end of the 17th century, with the proportion of *to*-infinitives still exceeding 17% in the period 1651–1700. In other words, the establishment of the use of bare infinitives is slightly slower in *EMEPS* than earlier studies suggest. Example (11) illustrates the use of *to*-infinitives in the second half of the 17th century:

(11)   This *made* the Great Hippocrat *to fall* out into this hyperbolick Expression.
(Sir Edward Eizat, 1695)

4.   Linguistic factors affecting the choice of *to-* and bare infinitives

4.1   Preliminary remarks

While the overall development of causative *make* in Early Modern English is clear and rather straightforward, there are some linguistic factors known to have affected the choice of *to-* or bare infinitives, making the path of the development more complex. Earlier studies discussing this issue are so abundant that it is virtually impossible to present a full survey in this paper, but I would like to mention some major works.[10] Mustanoja (1960: 522) and Fischer (1995, 1996 and 1997) point out the conceptual "closeness" conveyed by bare infinitives in general (i.e. not necessarily regarding causative *make* alone), although they use different terms.[11] Mustanoja (1960: 522) employs the term "intimacy", stating: "As for ME usage, the general principle is that when the relation between the finite verb and the infinitive is felt to be intimate, as in the case of auxiliaries like *can*, *may*, *must*, *shall*, and *will*, the plain infinitive [= bare infinitive] is used". Fischer, in her discussion of infinitival complements, uses the term "directness", saying: "zero infinitives [i.e. bare infinitives] indicate a 'direct' relationship between what is expressed in the matrix verb and the infinitival complement, and *to*-infinitives an 'indirect' one" (Fischer 1997: 111).

More recent studies specifically dealing with the complementation of causative verbs include Lowrey (2013), who develops Fischer's idea on the basis of his analysis of some Middle English texts and introduces the concept of agentivity as a key factor

---

10. The works mentioned in this section are all concerned with Middle English. This is not by accident, since a number of verbs, like causative *make*, governed both *to-* and bare infinitives in Middle English. It goes without saying that the choice of *to-* or bare infinitives is possible only when the two forms are available as grammatical options.

11. This is a widely accepted view about infinitives in general, but it is important to be aware that it is simply a matter of tendency and not a rule. For the different implications expressed by *to-* and bare infinitives in Present-day English, see Egan (2008: 272) among many others.

determining the choice of bare infinitives rather than *to*-infinitives. In my own work, I have also discussed the complementation of causative *make* (Iyeiri 2012), where I examine some linguistic factors which may affect the choice of *to*- and bare infinitives in later Middle English. In the remainder of this section, I will develop some major arguments in Iyeiri (2012) and examine if the choice of infinitival forms with causative *make* displays any continuity between Middle English and Early Modern English.[12]

4.2   Complexity in general

Iyeiri (2012) is much dependent on Rohdenburg's (1996: 151) Complexity Principle, which states: "In the case of more or less explicit grammatical options the more explicit one(s) will tend to be favored in cognitively more complex environments".[13] As Rohdenburg himself remarks, explicitness is a broad concept and can be attained in various ways, but in general, explicit options tend to be those with more coding material (Rohdenburg 1996: 151–152).[14] With regard to the complementation of causative *make*, *to*-infinitives, which have more coding material, are considered to be more explicit than bare infinitives.

In examining the applicability of this principle to causative *make* in later Middle English in my earlier work, I used the following working definition of complexity:

> Examples are complex when the object of causative *make* consists of three words or more. Examples are also complex when elements other than the object [i.e. object/causee] intervene between *make* and its complement.   (Iyeiri 2012: 66–67)

The present section uses the same working definition to analyze the data in Early Modern English. Hence, examples (12) and (13) are "complex": in (12), the object/causee consists of "three or more words", and in (13), an adverbial intervenes between *make* and the complement verb:

---

12. The present paper is in a way paired with Iyeiri (2012), where I discuss the complementation of causative *make* in later Middle English. These two papers combined are supposed to describe the major shift from *to*- to bare infinitives used as complements of causative *make* in the history of English.

13. Rohdenburg has published a number of papers discussing this principle; see also Rohdenburg (1995, 1998, and 2000).

14. The fact that the more explicit option tends to have more coding material can be explained within the framework of more general cognitive principles. See Givón (1991: 87), for example, whose Quantity Principle includes: "A larger chunk of information will be given a larger chunk of code". Although Givón's principles are designed mainly for the analysis of contemporary English, they are most likely applicable to the history of English.

(12)   This thynge maketh *me and many other* to meruayle. (Sir Thomas Elyot, 1541)

(13)   … which made *him shortly after* sing, Fortune my soe, &c.

(Sir Robert Naunton, 1641)

This definition yields 230 "complex" examples in the data of the present study, whose complementation patterns are summarized in Table 3:

Table 3. Complements of active causative *make* under "complex" circumstances in *EMEPS*

	To-infinitives	Bare infinitives	Totals
A16th–1 (1501–1550)	44 (88.0%)	6 (12.0%)	50
A16th–2 (1551–1600)	35 (58.3%)	25 (41.7%)	60
A17th–1 (1601–1650)	43 (57.3%)	32 (42.7%)	75
A17th–2 (1651–1700)	11 (24.4%)	34 (75.6%)	45

There is a clear tendency for "complex" examples to select *to*-infinitives at the beginning of the Early Modern English period: 88% of the relevant examples of active causative *make* are followed by *to*-infinitives in 1501–1550. On the other hand, the table also shows that bare infinitives become increasingly common even in the "complex" examples in Early Modern English, which is undoubtedly due to the overall rise of bare infinitives with active causative *make* in the history of English. It is, therefore, necessary to compare and contrast the figures in Table 3 with the overall situation of active causative *make*. This is graphically presented in Figure 1:
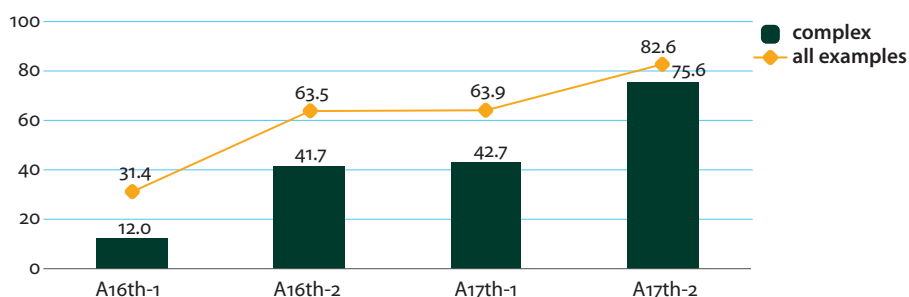
Figure 1. The proportions of bare infinitives in the "complex" examples and in the entire data of active causative *make* in *EMEPS* (%)

As Figure 1 demonstrates, the proportion of bare infinitives in the dataset of "complex" examples is invariably smaller than in the dataset comprising all examples, although the gap between the two datasets decreases over time. Hence, "complexity" is a functional factor in the choice of *to*- and bare infinitive as complements of causative *make* throughout the Early Modern English period. In other words,

to-infinitives, which are more "explicit" than bare infinitives, are more likely to be selected in "complex" environments than in other environments, though less so towards the end of the Early Modern English period, when the use of bare infinitives comes to be increasingly established with active causative *make*.

4.3   The object/causee of *make*

To examine the factor of complexity more closely, I will explore in this section the types of objects/causees of causative *make*. More specifically, I will investigate the choice of infinitival forms in the examples with personal-pronoun objects/causees and those with objects/causees of the other types with the assumption that personal pronouns are less complex than other object types such as noun phrases. This is also a development from my earlier work, where I demonstrate that, in later Middle English, *to*-infinitives are more likely to be selected when the object/causee of *make* is a noun phrase rather than a personal pronoun (see Iyeiri 2012: 65–66).

As a matter of fact, the preference for *to*-infinitives in noun-phrase environments seems to be a fairly strong tendency as it is noted by a number of previous studies dealing with different periods in the history of English: Ando (1976: 532–533), Rohdenburg (1995: 375–376, 1996: 155–159), and Rohdenburg & Schlüter (2000: 446–452) are just some of them. Among these, Rohdenburg (1995, 1996) and Rohdenburg & Schlüter (2000) use the Complexity Principle as an explanatory device on the presupposition that noun phrases are more complex than personal pronouns and therefore are more likely to select *to*-infinitives, which are a more explicit option than bare infinitives (see also Rohdenburg 1996: 152–155). The discussion below relies on the same framework.

The results obtained from *EMEPS* also corroborate these earlier studies. Tables 4 and 5 display the distribution of *to*- and bare infinitives with personal-pronoun objects/causees and objects/causees of the other types respectively.[15]

Table 4. *Make* + personal pronouns + infinitives in *EMEPS*

	To-infinitives	Bare infinitives	Totals
A16th–1 (1501–1550)	68 (59.1%)	47 (40.9%)	115
A16th–2 (1551–1600)	44 (23.2%)	146 (76.8%)	190
A17th–1 (1601–1650)	41 (25.0%)	123 (75.0%)	164
A17th–2 (1651–1700)	22 (11.3%)	172 (88.7%)	194

15. Examples that include both a personal pronoun and a noun phrase are excluded from the analysis. Example (12) quoted in the previous section is a case in point. Also excluded are those without any objects/causees at all. Hence, Tables 4 and 5 combined present fewer examples of active causative *make* than in Table 2 above.

Table 5. *Make* + other types of objects/causees + infinitives in *EMEPS*

	To-infinitives	Bare infinitives	Totals
A16th–1 (1501–1550)	81 (79.4%)	21 (20.6%)	102
A16th–2 (1551–1600)	68 (56.7%)	52 (43.3%)	120
A17th–1 (1601–1650)	71 (47.0%)	80 (53.0%)	151
A17th–2 (1651–1700)	37 (25.7%)	107 (74.3%)	144

Examples include:

(14)   Make *them* to vanishe lyke the smoke that they be no more sene.

(Roger Edward, 1570)

(15)   … that ye make *hym* runne   (Stephen Gardiner, 1546)

(16)   Neither do I make *S. Augustine* to speake false latine, but so as the bookes which I consulted, report him to speake.   (Harding Thomas, 1565)

(17)   for this also makes *the matter* smell further of then the poison could infect.

(Martin Aray, 1599)

Examples (14) and (15) illustrate the use of personal-pronoun objects/causees, while (16) and (17) exemplify objects/causees of the other types.

As Tables 4 and 5 elucidate, bare infinitives steadily increase from the first half of the 16th century onwards with both types of objects/causees. However, the proportion of bare infinitives is always higher when the object/causee of *make* is a personal pronoun (Table 4) than when it is a more complex noun phrase (Table 5). In the second half of the 17th century, for example, the proportion of bare infinitives reaches almost 90% when the object/causee is a personal pronoun (see Table 4), while the corresponding proportion for the object/causee of the other types stays around 75% (see Table 5). In other words, the expansion of bare infinitives with *make* is always more advanced when the object/causee is a personal pronoun.[16] It is also relevant to note that the proportion of bare infinitives in the context of "*make* + personal pronoun" (Table 4) is consistently higher than in the dataset comprising all attestations of causative *make* (Table 2). Conversely, bare infinitives in the context of "*make* + noun phrase" consistently take up a lower share than in the dataset in Table 2. Hence, the Complexity Principle is applicable to the data of this study as well.

---

16. The aim of the present section is to see the contrast in general between the sentences with personal-pronoun objects/causees and with objects/causees of the other types. Some examples have additional factors of complexity, but this applies equally to both categories.

4.4   Coordination of infinitives

The third factor to be considered is the coordination of infinitives in the complement of causative *make*, as in:

(18)   … which made him *to abide* in the Doctrine, and *follow* the Example of his
       Captain Christ Jesus …   (Francis Patchet, 1678)

In example (18), two infinitives following causative *make*, namely *to abide* and
follow, occur in coordination.[17]

     It has been argued in earlier studies that, in coordination, infinitive markers
are often reduced in the second complement (see, e.g., Svartvik & Quirk 1970: 402–
403): a sequence of the *to*-infinitive plus the bare infinitive is common, whereas
the opposite order, i.e. the bare infinitive followed by the *to*-infinitive, is not. This
propensity has been confirmed in my analysis of selected later Middle English
texts, where only two of a total of 43 relevant examples exhibit an increase of infinitival marking (i.e. from lighter to heavier marking) (Iyeiri 2012: 68–69).[18] At the
same time, however, 17 of the relevant examples employ the same infinitival form,
i.e. *to*-infinitives plus *to*-infinitives or bare infinitives plus bare infinitives (Iyeiri
2012: 69). It is, therefore, more accurate to state that infinitival markings are either
maintained or reduced in the second complement.

     This seems to be a fairly general tendency, which is applicable to different stages
of the history of English and consequently to Early Modern English as well. Of the
71 relevant examples in *EMEPS*, only two display an increase of marking, i.e. from
the bare infinitive to the *to*-infinitive:

(19)   But to make the Angels *garde* vs, and *to pitch round* about vs for our safetye, what
       an vnspeakeable prerogatiue is it of his incomparable mercye and bottomles
       compassion towardes the sheepe of his pasture?   (William Fisher, 1592)

(20)   But this ought not to make vs *disclayme or to neglect* the daily reading of Gods
       sacred Word.   (Francis Cartwright, 1621)

While (20) is syntactically fairly simple, in that the direct object *the daily reading
of Gods sacred Word* is shared by the two infinitives, i.e. *disclayme* and *to neglect*,

---

17. As pointed out earlier, this is the only section that deals with all infinitives occurring in
coordination, while in the other sections the first infinitive alone is considered in examples like
this.

18. In Iyeiri (2012), where I investigate causative *make* in Middle English, I use the following
scale: *that*-clauses (the heaviest marking) > *for to*-infinitives > *to*-infinitives > bare infinitives (the
lightest marking). In practice, however, the dataset does not yield any examples of the shift from
the *that*-clause to the infinitive.

example (19) is more demanding cognitively. It shows the repeated use of *vs*, but with different functions: the first instance of *vs* is a direct object, while the second appears in a prepositional phrase. The increase of the infinitival marking, namely from the bare infinitive to the *to*-infinitive, in (19) may be explicable from the perspective of the Complexity Principle, since the syntactic environment of the second complement is clearly more complex than the first one. It is important to be aware, however, that examples (19) and (20) are clearly exceptional in the entire data. Even in cognitively demanding circumstances, the general trend is for infinitives in coordinated structures to maintain their infinitival marking or shift from heavier to lighter marking. In the data under analysis, most examples illustrate a sequence of the *to*-infinitive plus *to*-infinitive(s) (15 examples in all), the bare infinitive plus bare infinitive(s) (37 examples in all), or the shift from the *to*-infinitive to bare infinitive(s) (17 examples in all). As expected, the combination of two bare infinitives is increasingly frequent as time passes, which is due to the historical expansion of bare infinitives in general. Of the 37 examples of "the bare infinitive plus bare infinitive(s)", 27 are attested in the 17th century, while the remaining 10 are found in the 16th century. And, of these ten 16th-century examples, 9 are attested in the latter half of the 16th century.

For the purpose of explaining various factors related to the choice of infinitival forms, the present study has so far relied on Rohdenburg's Complexity Principle. In this particular section, however, it may be appropriate to utilize more general principles concerning the relationship between information and coding. Givón's (1991: 87) Quantity Principle, for example, runs as follows:[19]

The Quantity Principle

a. A larger chunk of information will be given a larger chunk of code.
b. Less predictable information will be given more coding material.
c. More important information will be given more coding material.

Item (b) of this list is perhaps the most relevant to the discussion of infinitival coordination in the complement. The second element in a coordinate structure is more "predictable", at least in terms of its syntax and consequently in its relationship to the matrix verb, and this makes it unnecessary for the element to have heavier marking.[20] Hence, infinitival marking is likely to remain the same or to be reduced in the second (or later) complement.

---

19. The Quantity Principle is just one of the many tendencies presented by Givón (1991); see also footnote 14 above.

20. As this section focuses on the order of complements only and does not discuss the length of complements or the relative importance of the first and the second complements, the first and the third (i.e. (a) and (c)) of the Quantity Principle will not be considered here.

Supposing that higher syntactic predictability leads to the reduction of cognitive complexity, it may be possible to explain the tendency in coordinated infinitives by using Rohdenburg's Complexity Principle as well. The parsing of the second (or later) complement is easier or less complex, and therefore the bare infinitive, namely the less explicit form, is selected there. This is a matter of interpretation, though, since the existence of coordinated complements itself makes the sentence structurally and cognitively *more* complex. Also, the second (or later) complements are more distanced than the first one from the matrix verb *make*, and this may be interpreted as a case of complexity.

4.5   Infinitives of different verbs

Finally, I will discuss the relationship between the lexical item used as the infinitive and the choice of infinitival marking. Apparently, there is a noticeable tendency for some verbs to occur as *to*-infinitives and for others to occur as bare infinitives. I will start with the latter group, which include *appear* and *believe*. Table 6 shows the frequencies of *to appear* and *appear* in the complement of active causative *make*:

Table 6. *Appear* with active causative *make* in *EMEPS*

	To-infinitives (*to appear*)	Bare infinitives (*appear*)	Totals
A16th–1 (1501–1550)	1 (100%)	0 (0%)	1
A16th–2 (1551–1600)	2 (40.0%)	3 (60.0%)	5
A17th–1 (1601–1650)	5 (23.8%)	16 (76.2%)	21
A17th–2 (1651–1700)	0 (0%)	45 (100%)	45

Some illustrative examples are:

(21)   But I hope the eternall maiestie of God the sole disposer of all things will also *make* this *to appeare* in his good time.   (John Davis, 1595)

(22)   And soone he found in his first solitude it was but his shadow had *made* him *appeare* so mighty.   (William Habington, 1641)

Although the number of relevant examples in the data is rather limited, it is reasonable to infer from the above figures that the bare infinitive was already the dominant form of *appear* by the beginning of the 17th century or perhaps even earlier. Thus, *appear* seems to be ahead of other verbs regarding the shift to bare infinitives.

Incidentally, *appear* also tends to occur without *to*, even when causative *make* is used in the passive, as in (23):

(23)   … which in due time shall be made *appear* to the world

(Robert Norwood, 1652)

Although passive *make* is beyond the scope of this research, (23) is interesting, as it confirms that *appear* has a notably strong tendency to occur as a bare infinitive.

A similar, or even stronger tendency to occur as bare infinitives is observed with *believe*. Table 7 displays the occurrences of *to-* and bare infinitives when *believe* is in the complement of active causative *make*:

Table 7. *Believe* with active causative *make* in *EMEPS*

	To-infinitives (*to believe*)	Bare infinitives (*believe*)	Totals
A16th–1 (1501–1550)	1 (8.3%)	11 (91.7%)	12
A16th–2 (1551–1600)	0 (0%)	15 (100%)	15
A17th–1 (1601–1650)	0 (0%)	23 (100%)	23
A17th–2 (1651–1700)	0 (0%)	11 (100%)	11

The following are some illustrative examples:

(24)   Likewise Lawers, aduocates, sergeants, attorneis and procters are theues before God, when they for their owne gaynes do counsell a man to wage the lawe, *makynge* hym *to beleue* that hys mater is good, when in dede they thinke it naught.   (Thomas Cranmer, 1548)

(25)   Your Attention then, *made* me *believe* it not altogether Impertinent.

(Miles Barne, 1682)

As Table 7 demonstrates, bare infinitives are predominant with *believe* from the very beginning of the 16th century through to the end of the 17th century.

One possible explanation for the strong tendency for *appear* and *believe* to occur as bare infinitives is that both have an unstressed first syllable. Schlüter (2005) makes a thorough survey of various stress patterns in relation to the choice of grammatical forms in Modern English, including the choice of *to-* and bare infinitives with *make*, and shows that stress patterns affect the choice of the infinitival form: *to* is more likely to be followed by an initially stressed infinitive than by an initially unstressed one (2005: 185–209).[21] To return to *appear* and *believe*, the introduction of *to* before the infinitive would create an unwanted sequence of two unstressed syllables, which may have led to the preferred use of bare infinitives with these verbs.[22]

---

21. One is also reminded of Kiparsky's (1987: 194) reference to the metrical nature of language, which maintains that "language itself has the attributes which we associate with 'metrical' systems".

22. For the sake of contrast, it is relevant to mention that the verb *know*, another relatively frequent verb, which is monosyllabic and hence has stress on the first syllable, occurs in the entirety

By contrast, certain verbs are less likely to occur as bare infinitives, among which the verb *be* is one. Table 8 presents the frequencies of this verb used with or without *to* in the complement of *make*:

Table 8. *Be* with active causative *make* in *EMEPS*

	To-infinitives (*to be*)	Bare infinitives (*be*)	Totals
A16th–1 (1501–1550)	29 (90.6%)	3 (9.4%)	32
A16th–2 (1551–1600)	20 (87.0%)	3 (13.0%)	23
A17th–1 (1601–1650)	29 (100%)	0 (0%)	29
A17th–2 (1651–1700)	22 (88.0%)	3 (12.0%)	25

Some relevant examples are:

(26)   … since the blacke Mantelled night *makes* every thing *to be* silent
(Edward Reynolds, 1642)

(27)   It is thught they wil *make* you *be* cald vpon shortly to be Alderman of the Stilliard.
(Thomas Nash, 1592)

When *be* is in the complement, it occurs most frequently in the form *to be* rather than simple *be*, and this trend continues even till the end of the 17th century, when bare infinitives are otherwise much more commonly encountered. My previous work on later Middle English also shows that the use of *to be* was predominant as opposed to simple *be* in the 15th century: the proportion of *to*-infinitives was as large as 97.1% (Iyeiri 2012: 69–71).[23] When compared with this proportion, the use of bare infinitives has expanded slightly by the Early Modern English period (see Table 8), but not to a significant extent. The verb *be* still displays a strong tendency to occur with *to* in the Early Modern English period, when bare infinitives are increasingly dominant elsewhere.

It would be interesting to see further details of the relevant examples, namely whether *be* in the complement occurs in the passive infinitive construction, as in (28), or in other constructions like "(*to*) *be* + adjective/noun phrase", as in (29):

(28)   & that they wolde make me *be rewarded* by some of them …
(Henry Watson, 1510)

---

of the corpus 7 times as a *to*-infinitive as opposed to 19 times as a bare infinitive. The ratio of bare infinitives with *know* (73.1%) is clearly smaller than that of bare infinitives with *appear* and *believe*, which may be attributed to the difference of the stress patterns between *know* and the other two verbs. It is tempting to apply this analysis to all of the verbs attested in *EMEPS*. Unfortunately, however, most verbs provide only about 10 or even fewer examples in the entire data, and do not provide statistically significant results.

23. The analysis is based on 7 works printed by William Caxton.

(29)   … since the blacke Mantelled night makes every thing *to be silent* …

(Edward Reynolds, 1642)

Unfortunately, however, a further division of the data in Table 8 would not be reveal-
ing, as there are only 3 examples of bare infinitives in each of the periods A16th–1,
A16th–2, and A17th–2, and no examples of bare infinitives in A17th–1. All I can
say here is that about half of the examples in the entire table illustrate the "(*to*) *be* +
past participle" (passive infinitive) construction.

In terms of the Complexity Principle, passivity is considered to be conceptually
more complex. It can, therefore, invite the use of *to*-infinitives, which are consid-
ered to be more explicit than bare infinitives. To prove this, however, one needs
to conduct research on a larger scale, since those examples which are not in the
passive infinitival construction also show the dominance of *to*-infinitives, making
the difference between passive infinitives and non-passive ones rather obscure in
a small set of data (see Table 8). Fischer's (1995: 8) statement about late Middle
English in general (not necessarily concerning the causative) that "the *to*-infinitive
is the rule when the infinitive or the matrix verb is in the passive form" is interesting
in this relation,[24] but it is hasty at this stage to view the data available as supporting
evidence. I can say, however, that the use of *to*-infinitives is no longer "the rule" in
the Early Modern English period, when bare infinitives are increasingly established,
though infinitives with *to* are indeed favored by *be*.

5.   Conclusion

To summarize the above discussion, the use of bare infinitives is not at all estab-
lished in the first half of the 16th century. However, this situation is reversed by the
second half of the 17th century, when more than 80% of the relevant examples in
the active voice yield bare infinitives.

Moreover, some linguistic factors seem to affect the choice of *to*-infinitives vs.
bare infinitives with causative *make*. First of all, there is a fairly clear correlation be-
tween the complexity of sentences in general and the selection of infinitival forms:
complex sentences have a fairly clear tendency to select *to*-infinitives instead of bare
infinitives. Furthermore, it has been shown above that with regard to the object/

---

24. Fischer's remark is based on her investigation of Chaucer and the *Paston Letters*. As it stands,
her comment refers to the passivity of not only the complement but also the matrix verb. The
present study is, however, concerned with the complement only. As mentioned above, the passive
use of *make* itself is outside the ambit of this study.

causee of *make*, personal-pronoun objects/causees are more likely to be associated with bare infinitives than objects/causees of the other types. This is also explicable within the framework of the Complexity Principle.

The two other factors discussed above are the coordination of infinitives and the lexical item used in the complement. The analysis of the former factor has shown that, in the coordination of infinitives, infinitival markings are likely to stay the same or to be reduced in the second and later complements. Thus, bare infinitives followed by *to*-infinitives are unusual, while *to*-infinitives followed by *to*-infinitives or bare infinitives and bare infinitives followed by bare infinitives are common. Over time, the frequency of successive occurrences of bare infinitives increases, which is simply due to the overall expansion of bare infinitives with active causative *make* in the Early Modern English period.

As for the relationship between the choice of infinitival forms and the lexical item used in the complement, *appear* and *believe* are commonly found as bare infinitives, whereas *be* tends to occur as *to be*. As for *appear* and *believe*, the preference for the bare form may be related to their stress patterns, as both have an unstressed syllable in initial position – the addition of *to* would result in an undesirable sequence of two unstressed syllables. With regard to the verb *be,* on the other hand, it is the *to*-infinitival form that is visibly most frequent. It is interesting that a number of instances of *to be* are observed in the passive infinitive, although accounting for the relationship between the passivity and the choice of *to be* in Early Modern English requires further investigation.

All in all, it has been seen that a number of linguistic factors still played a role in the choice of *to*- and bare infinitives in Early Modern English. In due course, however, they ceased to be operative, as the use of bare infinitives became increasingly established with causative *make* in Modern English.

Acknowledgements

References

Andersson, Evert. 1985. *On Verb Complementation in Written English*. Malmö: CWK Gleerup.

Ando, Sadao. 1976. *A Descriptive Syntax of Christopher Marlowe's Language*. Tokyo: University of Tokyo Press.

Egan, Thomas. 2008. *Non-finite Complementation: A Usage-based Study of Infinitive and -ing Clauses in English*. Amsterdam: Rodopi. doi:10.1163/9789401205542

Fanego, Teresa. 1994. Infinitive marking in Early Modern English. In *English Historical Linguistics 1992: Papers from the 7th International Conference on English Historical Linguistics, Valencia, 22–26 September 1992* [Current Issues in Linguistic Theory 113], Francisco Fernández, Miguel Fuster & Juan José Calvo (eds), 191–203. Amsterdam: John Benjamins. doi:10.1075/cilt.113.20fan

Fischer, Olga. 1995. The distinction between *to* and bare infinitival complements in Late Middle English. *Diachronica* 12: 1–30. doi:10.1075/dia.12.1.02fis

Fischer, Olga. 1996. Verbal complementation in early ME: How do the infinitives fit in? In *English Historical Linguistics 1994: Papers from the 8th International Conference on English Historical Linguistics (8 ICEHL, Edinburgh, 19–23 September 1994)* [Current Issues in Linguistic Theory 135], Derek Britton (ed.), 247–270. Amsterdam: John Benjamins. doi:10.1075/cilt.135.18fis

Fischer, Olga. 1997. Infinitive marking in Late Middle English: Transitivity and changes in the English system of case. In *Studies in Middle English Linguistics*, Jacek Fisiak (ed.), 109–134. Berlin: Mouton de Gruyter. doi:10.1515/9783110814194.109

Givón, Talmy. 1991. Isomorphism in the grammatical code: Cognitive and biological consideration. *Studies in Language* 15: 85–114. doi:10.1075/sl.15.1.04giv

Iyeiri, Yoko. 2010. *Verbs of Implicit Negation and Their Complements in the History of English*. Amsterdam: John Benjamins & Tokyo: Yushodo Press. doi:10.1075/z.155

Iyeiri, Yoko. 2011. Early Modern English prose selections: Directions in historical corpus linguistics. *Memoirs of the Faculty of Letters, Kyoto University* 50: 133–199.

Iyeiri, Yoko. 2012. The complements of causative *make* in Late Middle English. In *Middle and Modern English Corpus Linguistics: A Multi-dimensional Approach* [Studies in Corpus Linguistics 50], Manfred Markus, Yoko Iyeiri, Reinhard Heuberger & Emil Chamson (eds), 59–73. Amsterdam: John Benjamins. doi:10.1075/scl.50.08iye

Kiparsky, Paul. 1987. On theory and interpretation. In *The Linguistics of Writing: Arguments between Language and Literature*, Nigel Fabb, Derek Attridge, Alan Durant & Colin MacCabe (eds), 185–198. Manchester: Manchester University Press.

Los, Bettelou. 2005. *The Rise of the* To-*Infinitive*. Oxford: OUP. doi:10.1093/acprof:oso/9780199274765.001.0001

Lowrey, Brian. 2013. Syntactic and semantic variation: Middle English causatives and complement types. In *Periphrasis, Replacement and Renewal: Studies in English Historical Linguistics*, Irén Hegedűs & Dóra Pődör (eds), 92–124. Newcastle upon Tyne: Cambridge Scholars.

Mustanoja, Tauno F. 1960. *A Middle English Syntax, Part 1: Parts of Speech*. Helsinki: Société Néophilologique.

Rohdenburg, Günter. 1995. On the replacement of finite complement clauses by infinitives in English. *English Studies* 76: 367–388. doi:10.1080/00138389508598980

Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7: 149–182. doi:10.1515/cogl.1996.7.2.149

Rohdenburg, Günter. 1998. Syntactic complexity and the variable use of *to be* in 16th to 18th century English. *Arbeiten aus Anglistik und Amerikanistik* 23: 199–228.

Rohdenburg, Günter. 2000. The complexity principle as a factor determining grammatical variation and change in English. In *Language Use, Language Acquisition and Language History: (Mostly) Empirical Studies in Honour of Rüdiger Zimmermann*, Ingo Plag & Klaus Peter Schneider (eds), 25–44. Trier: Wissenschaftlicher Verlag.

Rohdenburg, Günter & Schlüter, Julia. 2000. Determinanten grammatischer Variation im Früh- und Spätneuenglischen. *Sprachwissenschaft* 25: 443–496.

Sawada, Mayumi. 1997. Causative verbs in Chaucer. *ERA* (The English Research Association of Hiroshima) 15: 1–16.

Schlüter, Julia. 2005. *Rhythmic Grammar: The Influence of Rhythm on Grammatical Variation and Change in English*. Berlin: Mouton de Gruyter. doi:10.1515/9783110219265

Svartvik, Jan & Quirk, Randolph. 1970. Types and uses of non-finite clause in Chaucer. *English Studies* 51: 393–411. doi:10.1080/00138387008597391

Sugiyama, Ryuichi. 1988. Jugo seiki sanbun no shieki kobun niokeru futeishi hyoshiki nitsuite (On the marker of the infinitive governed by causative verbs in fifteenth-century prose). *Studies in Medieval English Language and Literature* 3: 43–55.

Semantic and lexical shifts with the "*into*-causative" construction in American English

Mark Davies and Jong-Bok Kim

Brigham Young University / Kyung-Hee University

In this paper, we consider several lexical and semantic shifts with the "*into*-causative" construction (e.g. *Sue talked them into leaving*) in American English since the early 1800s. The study is based on more than 11,000 tokens (including 680 different matrix verbs) in several large corpora, including *COHA*, *COCA*, *TIME*, and *GloWbE*. We consider overall changes in the semantic classes of verbs that can be used in the construction (e.g. verbs of force, persuasion, trickery). We then look in some detail at changes with "neutral" verbs (e.g. *lead*) and "positive" verbs (e.g. *encourage*), "indirect causation" and a hybrid construction involving the *way* construction.

Keywords: corpus, causative, into, American English, lexical, semantic, diachronic

1. Introduction

The typical "*into*-causative" construction is composed of a verb of force or persuasion (*beguile, fool*, and *embarrass* below) + *into* + an *-ing*-clause:

(1) a. They would *beguile* us *into believing* that we are to fall down and worship the image.[1] (*Corpus of Historical American English (COHA)*, NF, *1847*)
   b. I don't see why you should *fool* yourself *into thinking* you're sorry.
      (*COHA*, FIC, *1922*)
   c. we almost *embarrass* him *into going* along if we have to
      (*COHA*, NEWS, *1986*)

---

1. The abbreviations of the genre categories in *COHA* are as follows: FIC = Fiction, NEWS = Newspapers, NF = Non-fiction.

The "*into*-causative" construction in Present-day English has been studied in some detail, in large part because of the interesting ways in which meaning and structure are related. What has been much less studied, however, is the historical development of the construction, which will be the focus of the present study. Without understanding where the "*into*-causative" has come from and how it has been changing during the past 100–200 years, it is difficult to fully understand the construction in Present-day English.

In this study, we will attempt to answer a number of questions about changes in the lexical and semantic properties of the "*into*-causative" construction, including the following:

Lexical (and frequency)
– Has the construction become more common or less common in English over time? Also, how much lexical diversity has there been over time, in terms of new matrix verbs (e.g. *talk, coerce, trick*) that have been used in different periods? (Section 4)

Semantic
– What changes have there been in the semantic classes of verbs that are used in the construction (e.g. verbs of force, persuasion, trickery, etc)? (Section 5)
– Is the use with "neutral" verbs (e.g. *lead*) and "positive" verbs (e.g. *encourage*) a recent innovation (as some have suggested), or have these uses been around for a much longer period of time? (Sections 6–7)
– What changes have there been in the degree of "directness" of force or persuasion by X on Y to do Z (e.g *John* (X) *coerced Mary* (Y) *into coming* (Z))? (Section 8)
– How has the "*into*-causative" construction been related over time to the *way* construction (e.g. *They lied their way into taking our country into war*)? (Section 9)

And finally (in the concluding Section 10), what does an in-depth study of this construction tell us about the value of large and diverse historical corpora, and their role in researching syntactic and semantic change?

2. Previous research

As has been mentioned, the "*into*-causative" construction in contemporary English has been the focus of quite a number of studies (see Bridgeman et al. 1965; Rudanko 1991, 2002, 2005, 2006, 2011, 2015a, 2015b; Francis et al. 1996; Hunston & Francis 2000; Gries & Stefanowitsch 2003; Wulff et al. 2007; Kim & Davies 2016). As we have noted, however, the historical development of the construction has received much less attention. And as we will see, this is due to the lack (until recently) of useful and robust historical corpora.

The articles and chapters by Rudanko (2000, 2005, 2006, 2011, 2015a, 2015b) comprise virtually all that has been published about the history of the "*into*-causative" construction. Rudanko (2000: Chapter 5) examines the construction in the 1700s and in the late 1900s (1970s and later). The data for the 1700s is based on the *Chadwyck-Healey Corpus of Eighteenth-Century Fiction* (which later became part of the *Literature Online* database),[2] and the data from the late 1900s comes from the *Bank of English*[3] and the *British National Corpus*.[4] Rudanko finds that in the 1700s, the construction is still quite rare. In a corpus comprising about 10 million words, he finds a total of just 28 tokens (a normalized frequency of 2.8 per million words), with 19 different matrix verbs. For each verb he gives the relevant tokens from the corpus, and he provides some useful semantic categorization of the verbs as well. In terms of changes from the 1700s to the late 1900s, he suggests that there is a much wider range of verbs that allow the construction in the later period, and that the "*into*-causative" construction is best analyzed from a Construction Grammar perspective (see, e.g., Goldberg 1995). Rudanko (2011) is based primarily on data from contemporary English, but Chapter 2 of the book does deal with "*into*-causative" in the *Corpus of English Novels* (*CEN*).[5] This corpus contains 25 million words of text from publicly available novels (mainly from Project Gutenberg) from 1880–1922, including 12.3 million words from British novels and 5.9 million words from American novels. In total, Rudanko finds 93 tokens (7.6 per million words) of the construction with 50 different matrix verbs in the British novels and 57 tokens (9.7 per million words) with 32 different matrix verbs in the American novels. In terms of semantic categorization, he suggests that verbs of deception (e.g. *betray, cheat, deceive, entrap, inveigle, mislead*, and *trap*) are the most common type of verb, and he suggests that verbs of force and pressure were more common in British than in American novels. Finally, he compares the data from 1880–1922 with two smaller corpora of novels from the late 1900s, the 5.5 million-word *British Books Corpus* and the 5.6 million-word *United States Books Corpus* (both available from the *Bank of English*), and he shows that the construction has increased in frequency over time: there are 97 tokens (17.6 per million words) with 53 types in the *British Books Corpus* and 80 tokens (14.3 per million words) with 40 types in the *United States Books Corpus*. In this same chapter, Rudanko also comments briefly on the 66 tokens of the construction in the Brown family of corpora – one million words each from US 1961 (*Brown*), US 1991 (*Frown*), UK 1961 (*LOB*), and UK 1991 (*FLOB*).[6] On the basis of

---

2. See <http://collections.chadwyck.co.uk/marketing/home_c18f.jsp>

3. See <http://www.collins.co.uk/page/The+Collins+Corpus>

4. See <http://www.natcorp.ox.ac.uk/>

5. See <https://perswww.kuleuven.be/~u0044428/>

6. See <http://www.helsinki.fi/varieng/CoRD/corpora/>

the distribution of these 66 tokens over the four corpora (*Brown*: 18, *Frown*: 21, *LOB*: 6, *FLOB*: 21), he notes that there has been a likely increase in the overall frequency of the construction in the last 30–40 years, but suggests that we need much larger corpora to accurately describe the historical changes. Finally, Rudanko (2005) and Rudanko (2006) are both based on texts from just contemporary English, although there is an attempt to extrapolate back historically from this data. Rudanko (2005) suggests that there has recently been an increase in "neutral" verbs (e.g. *lead, impel, induce, influence, motivate, prompt*, and *stimulate*; see Section 6 below) and he suggests that this is more common in British than in American English. Rudanko (2006) provides data for more than fifty matrix verbs that had not been mentioned in previous studies, and he suggests that British English has been more innovative in moving from [*to* V] to [V-*ing*] complements with some of these innovative verbs (cf. Rudanko 2000; Mair 2002; Vosberg 2003; Rohdenburg 2007).

3. Corpus data

Rudanko's studies represent an exceptionally careful and thorough analysis of the data that was available to him at the time. In each article, however, he laments the unavailability of larger and more systematic historical corpora of English on which to base the analyses. Our study is designed to address this issue. It is explicitly based on large and robust corpora that were not available to Rudanko as he wrote his insightful articles, but which have become available in the last 4 to 5 years.

The majority of the data for our study are taken from the following corpora of American English:

Table 1. Corpora used in this study

Corpus	Tokens	Corpus size	Texts	Period
Corpus of Historical American English (*COHA*)	3,874	400 million	100,000+	1810–2009
TIME Magazine Corpus of American English (*TIME*)	1,718	100 million	275,000+	1926–2006
Corpus of Contemporary American English (*COCA*)	5,848	450 million	190,000+	1990–2012
TOTAL	**11,440**	**950 million**	**565,000+**	

As can be seen, our study is based on a much larger data set of tokens than has been available for any previous study. It is based on more than 5,500 tokens with over 500 different matrix verbs for the 1810s–1980s (from *COHA* and *TIME*), and then another 5,800 or so tokens for the 1990–2000s (from *COCA*). This means that we

have about 100 times as much data as in a small set of corpora such as the Brown family of corpora (1960s–1990s), and 30–40 times as much data as from the *Corpus of English Novels* (1880–1922) that was used in Rudanko (2011). In addition, we have data from each decade during the past 200 years; this allows us to follow changes in the construction decade by decade, something which has not been possible in previous studies.

   In terms of the composition of the corpora, *COHA* is the largest structured corpus of historical English that is based on a variety of genres, and which maintains the same genre balance from year to year (see Davies 2012a, 2012b). It is more than 100 times as large as any other structured and balanced corpus of historical English, and it contains 207 million words of fiction (novels, short stories, plays and movie scripts), 97 million words from popular magazines, 40 million words from newspapers, and 61 million words from non-fiction books. The *TIME Magazine Corpus* is based on more than 275,000 articles in *TIME Magazine* from the 1920s to the 2000s. And finally, *COCA* is based on spoken, fiction, magazine, newspaper, and academic texts from 1990 to the current time (the last update is from December 2015, and it continues to be updated), and it maintains the same genre (and sub-genre) balance from year to year (see Davies 2009, 2010). Together, these corpora (all of which are freely available from <http://corpus.byu.edu>) provide a wealth of data on the "*into*-causative" construction, which greatly adds to the data from smaller corpora that have been presented in previous studies.

   The search query we have adopted is a simple one:

   (2)   [vv*] {0,4} into [v?g*]

This query searches for any string that is composed of a lexical verb followed by the preposition *into* and a verb ending in *-ing*, where the distance between the verb and *into* can be from zero to four. (In the *COHA* web interface, this would be into [v?g*] in the WORD(S) field, [vv*] in the COLLOCATES field, and a collocates span of 4 words left and 0 words right). Strings with a distance of 1–4 would find, for example:

   (3)   He talked { them$_1$ / the$_1$ men$_2$ / the$_1$ young$_2$ men$_3$ } into$_4$ coming.[7]

The context {0,4} represents 4 or less (including zero) collocate distances between the main verb and the into clause. The distance zero is to include examples like the following passive construction:

---

7.   It would not, however, find *he talked* {*the really old and crazy and nearly toothless*} *men into coming.* While the corpus web interface can find strings as long as 21 words in length, we found that the vast majority of all relevant tokens are found within this four word frame. For example, of 200 random tokens of [*talk + into* V-*ing*] in the corpus, 197 of them had the matrix verb within four words of *into*.

(4) a. which is a venture that I never was *persuaded into undertaking* before
(*COHA*, FIC, 1869)
  b. if I *was misled into making* statements that were untrue in my last meeting
(*COHA*, FIC, 1922)

4. Lexical diversity

Before looking at the increasing range of matrix verbs that participate in the construction, we will briefly consider the overall increase in the frequency of the "*into*-causative" construction over time in American English. The chart in Figure 1 shows the normalized frequency (per million words) of the "*into*-causative" in *COHA* in thirty-year segments from the early 1800s to the current time. As can be seen, the overall increase (with all verbs) has been quite consistent during the past 200 years, which supports the data found in Rudanko (2011: Chapter 2).
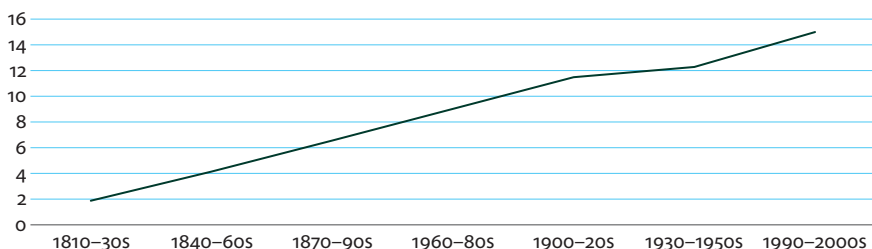
Figure 1. Overall increase in frequency (per million words) of the "*into*-causative" in *COHA*

An alternative view of the data is in Figure 2. This is taken from the *COHA* web interface itself, and it shows the normalized frequency by decade (1810s–2000s) of [lexical verb + pronoun + *into* + V-*ing*] (e.g. *talk them into coming*). While this particular string represents only a subset of all tokens of the "*into*-causative" construction (evidenced by the lower normalized frequency in each decade), it nevertheless shows that *COHA* provides clear evidence of the overall increase in frequency.

SECTION	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
FREQ	0	4	13	22	19	33	53	51	46	61	99	118	105	122	150	106	122	122	181	234
PER MIL	0.00	0.58	0.94	1.37	1.15	1.93	2.86	2.51	2.23	2.76	4.36	4.60	4.27	5.01	6.11	4.42	5.12	4.82	6.48	7.91

Figure 2. Overall increase in frequency of a subset of the "*into*-causative" (taken from the *COHA* web interface)

In addition to the overall increase in frequency, there has also been an increase in the number of matrix verbs that participate in the construction, and this is shown in Table 2. This table shows the overall normalized frequency in each of the seven periods (corresponding to Figure 1 above), followed by the number of new verbs in that period that occur at least once ($f \geq 1$) and twice ($f \geq 2$), as well as a list of the "new" verbs (i.e. which do not appear in previous decades) that occur at least twice (with the number of tokens of each verb in that period).

Table 2. Most frequent new verbs by time period in *COHA*

Period	Freq	$f \geq 1$	$f \geq 2$	New verbs (frequency ≥ 2)
1840s–1860s	4.13	67	25	frighten 10, deceive 5, persuade 5, force 5, cheat 4, delude 3, inveigle 3, irritate 3, seduce 3, shame 3, startle 3, surprise 3, talk 3, trick 3, provoke 2, quicken 2, tease 2, terrify 2, lure 2, draw 2, entice 2, flog 2, construe 2, bully 2, cajole 2
1870s–1890s	6.54	56	16	coerce 9, flatter 4, trap 4, fool 4, rouse 3, spur 2, stimulate 2, train 2, transform 2, gull 2, hoodwink 2, whip 2, worry 2, cow 2, crush 2, argue 2
1900s–1920s	9.01	92	20	bluff 8, hypnotize 7, kid 5, nag 4, throw 4, shock 3, bamboozle 3, starve 3, stampede 2, trip 2, push 2, anger 2, astonish 2, jar 2, jeer 2, jolly 2, browbeat 2, conjure 2, dupe 2, enchant 2
1930s–1950s	11.49	99	22	pressure 9, prod 7, maneuver 6, rush 5, threaten 3, needle 3, josh 3, spur 3, squeeze 2, steer 2, surprise 2, rationalize 2, panic 2, plunge 2, blackjack 2, con 2, corner 2, cow 2, interest 2, jockey 2, shake 2
1960s–1980s	12.28	119	13	embarrass 4, guide 3, kick 3, lock 2, nudge 2, sucker 2, taunt 2, terrorize 2, trigger 2, harry 2, change 2, co-opt 2, drag 2
1990s–2000s	14.99	93	15	throw 7, draft 3, move 3, whip 3, will 2, transform 2, chase 2, hook 2, invest 2, jolt 2, nurture 2, sink 2, steer 2, stun 2, sway 2
TOTAL		544		

Although *COHA* contains tokens for the "*into*-causative" construction with 544 different matrix verbs, this is in no way a comprehensive list of all possible verbs. In addition, there are 89 other verbs (beyond the 544 in *COHA*) that occur in the 100 million word *TIME* Corpus, but which are not in *COHA*. These include verbs such as *heffle, hornswoggle* (1920s); *egg, sting, catspaw* (1930s); *preach, pound, smooch* (1940s); *cramp, dope, jar* (1950s); *jolly, gig, stiffen* (1960s); *drill, hook, Svengali* (1970s); *blarney, blow* (1980s); and *euchre, sober* (1990). Examples are in (5).

(5)   a.   Last week Senator Heflin tried, at length, to *heffle* the Senate *into adopting* a resolution condemning the nameless bottle-thrower.
      (*TIME*, 1920s) [note the play on the senator's name]

      b.   This was supposed to have *stung* Dictator Stalin *into assuming* a defiant attitude.   (*TIME*,1930s)

      c.   Then his boss *catspawed* him *into marrying* a European mistress who was getting troublesome.   (*TIME*,1940s)

      d.   they had encouraged his wife to *smooch* the customers *into buying* more drinks   (*TIME*, 1940s)

      e.   he *Svengalied* willing authors *into writing* potboilers and racy romans clef   (*TIME*, 1970s)

      f.   Moynihan, who … *blarneyed* Nixon *into endorsing* the idea   (*TIME*, 1980s) [note: Moynihan was Irish]

Note the incredible range of lexical diversity with these verbs. And lest one think – even with these additional verbs – that the full range of possible verbs has been exhausted, there are another 47 verbs from *COCA* (for just the 1990s–2000s) that are not found in *COHA* or *TIME*. The most common of these (occurring three times or more in *COCA*) are *socialize, funnel, prompt, twist, whisk, tip, gain, game, group, grow, assemble, buffalo, distract, fashion, spoon, steamroll, jawbone*, and *misdirect*.

Indeed, the range of verbs is seemingly endless. As Hunston & Francis (2000: 103) ask (referring to the "into V-*ing* construction"), "are there any limits to the creativity of speakers: can we state categorically that something cannot be said?" Given the tens of thousands of lexical verbs in English, and the possibility that thousands of these could potentially be used in a causative sense (see Goldberg 1995), it is undoubtedly the case that if we had another billion or so words of corpus data (in addition to the 950 million words in *COHA*, *TIME*, and *COCA*), we would find many more verbs as well. In spite of this dizzying range of lexical diversity, we will see in Sections 5–7 that there are some interesting tendencies in terms of changes in semantic classes of verbs over time.

5.   Semantic changes: Some generalizations

With more than 680 different matrix verbs, it may seem a bit overwhelming to understand the principal changes in semantic categories of these verbs over time. But we can focus on general trends involving the most frequent verbs, which will account for a large percentage of all tokens. For example, the twenty verbs listed in Table 3 account for 1,935 of the total 3,874 tokens in *COHA*, or about 50% of all tokens.

Table 3. Most frequent verbs taking the "*into*-causative" (overall in *COHA*, all periods)

		TOT	1810–30s	1840–60s	1870–90s	1900–20s	1930–50s	1960–80s	1990–2000s
1	*talk*	415	0	3	6	13	104	141	148
2	*trick*	204	0	3	4	28	45	48	76
3	*bring*	117	5	11	21	20	28	24	8
4	*force*	117	0	5	11	30	21	37	13
5	*fool*	114	1	0	4	26	23	20	40
6	*deceive*	97	0	5	17	31	20	13	11
7	*frighten*	82	0	10	23	20	15	10	4
8	*coax*	81	2	5	13	7	7	19	28
9	*call*	71	11	20	21	12	5	2	0
10	*delude*	71	0	3	11	19	13	17	8
11	*coerce*	67	0	0	9	9	21	12	16
12	*mislead*	67	0	1	11	18	15	16	6
13	*scare*	67	1	1	9	10	17	9	20
14	*pressure*	57	0	0	0	0	9	27	21
15	*beguile*	57	1	12	19	11	10	3	1
16	*lure*	55	0	2	2	6	13	17	15
17	*lead*	53	1	4	15	9	11	7	6
18	*bully*	51	0	2	4	15	8	8	14
19	*goad*	48	0	0	1	10	12	12	13
20	*provoke*	44	0	2	2	5	8	14	13
	ALL	1935	22	89	203	299	405	456	461

In terms of overall changes in semantic classes, perhaps the most useful view of the data would be to look at the most frequent verbs in different time periods (see Rudanko 2011 for a similar approach, based on the smaller *CEN Corpus*). Table 4 identifies the top 15–20 verbs in *COHA* in each thirty-year time period by semantic class. The verbs *call* and *bring* are bolded and underlined; verbs of force are underlined (e.g. *force, coerce, drive, pressure*); verbs of persuasion are in italics (e.g. *coax, persuade, wheedle*); verbs of emotional effect are in bold (e.g. *surprise, startle, scare*); the "neutral" verb *lead* is in small caps; the one very frequent verb *talk* is in Verdana typeface; and all other verbs (primarily verbs of deception) are in normal font (e.g. *deceive, cheat, trick*).

It is important to note that not all verbs neatly fit into just one semantic class. For example, *goad* and *prod* are in a sense both verbs of force and of (strong) persuasion. There is no "iron-clad" semantic categorization here that would not be open to criticism. Nevertheless, we have used essentially the same semantic categories as Rudanko (2011), which in turn is based (largely) on the semantic categories in Hunston & Francis (2000), with the addition of categories such as "neutral".

Table 4. Most frequent verbs taking the "*into*-causative" in COHA, by semantic category (by thirty-year period)

1810–30s		1840–60s		1870–90s		1900–20s		1930–50s		1960–80s		1990–2000s	
call	11	call	20	frighten	23	deceive	31	talk	104	talk	141	talk	148
bring	5	beguile	12	call	21	force	30	trick	45	trick	48	trick	76
resolve	2	bring	11	bring	21	trick	28	bring	28	force	37	fool	40
coax	2	frighten	10	beguile	19	fool	26	fool	23	pressure	27	coax	28
		force	5	deceive	17	bring	20	coerce	21	bring	24	pressure	21
		betray	5	LEAD	15	frighten	20	force	21	fool	20	scare	20
		deceive	5	coax	13	delude	19	deceive	20	coax	19	coerce	16
		coax	5	persuade	13	mislead	18	scare	17	delude	17	lure	15
		persuade	5	mislead	11	betray	17	trap	17	lure	17	bully	14
		cheat	4	force	11	wheedle	16	drive	15	mislead	16	con	14
		LEAD	4	delude	11	bully	15	frighten	15	provoke	14	force	13
		delude	3	tempt	11	talk	13	argue	15	deceive	13	goad	13
		drive	3	coerce	9	call	12	mislead	15	coerce	12	provoke	13
		talk	3	drive	9	drive	11	blackmail	13	prod	12	seduce	13
		surprise	3	scare	9	beguile	11	delude	13	goad	12	deceive	11
		starve	3	inveigle	8	cajole	10	kid	13	trap	11	lull	10
		startle	3	betray	8	goad	10	lure	13	frighten	10	shock	9
		shame	3	wheedle	7	inveigle	10	goad	12	scare	9	bring	8
		seduce	3	entrap	6	scare	10	lead	11	bully	8	delude	8
		irritate	3	talk	6	shame	10	beguile	10	tease	8	shame	8
		inveigle	3							tempt	8		
		trick	3										

This data allows us to make a number of general observations. First, note the high frequency of *call* and *bring* in the early decades.[8] Second, the verb *talk* (technically a member of the semantic class "persuasion") has increased in frequency since the early 1900s, and has remained dominant since the 1930s. Third, by the 1840–1860s, nearly all of the semantic classes are already present: *call* and *bring*, *talk* (= persuasion), other persuasion verbs (e.g. *coax, persuade*), the neutral verb *lead*, verbs of emotional effect (e.g. *frighten, surprise*), verbs of force (e.g. *force, drive*), and verbs of deception (e.g. *beguile, deceive*). Fourth, in the 1930s–1950s there are more verbs of persuasion, whereas in the 1960s–1980s and 1990s–2000s there are more verbs of force. These frequency figures may point to the beginning of differing trends in verbs of persuasion and force; they would thus nicely tie in with the claim in Wulff et al. (2007) that British English uses verbs of force more than American English, whereas American English uses verbs of persuasion more than British English.

---

8. The vast majority of the tokens with *bring* and *call* are in the fixed phrase *call / bring something into being*, where *being* is probably best considered a noun.

Note that this assumes that we can accurately distinguish between the two classes of verbs. As we pointed out above, an "iron-clad" semantic distinction between verbs of persuasion and verbs of force will always be open to criticism

6. Neutral verbs

Rudanko (2005) suggests – based on data from Present-day British and American texts – that there has recently been a semantic shift in which "neutral" verbs have begun to be used with the "*into*-causative" construction. Rudanko defines neutral verbs as verbs in which there is "unflavored" (i.e. not necessarily negative or deceptive) interaction between the subject of the main clause and the subordinate clause, such as with the verbs *influence, impel*, *induce*, *lead*, *motivate*, *prompt*, and *stimulate*. The underlying assumption is that in earlier periods most (or all) causative verbs had negative prosody (e.g. *force, deceive, beguile*), and therefore the presence of these "neutral" verbs in Present-day English must somehow be explained. Because our study is the first one to provide robust data for these earlier periods, we can test this assumption.

The data from *COHA* shows that there in fact has been little or no increase with "neutral" verbs over time. Table 5 shows the number of tokens with these verbs in each of the seven time periods since the early 1800s. Note that in this case, we modified the search slightly to include all tokens where there were 1 to 3 words between the matrix and the embedded verb (cf. the formula shown in (2) above.)

Table 5. Frequency distribution of "neutral" verbs in *COHA*, by thirty-year period

	1810–30s	1840–60s	1870–90s	1900–20s	1930–50s	1960–80s	1990–2000s
lead	1	4	15	9	11	7	6
draw	0	2	3	3	2	10	2
guide	0	1	0	0	0	3	1
steer	0	0	0	0	2	0	2
stimulate	0	0	2	0	0	1	1
trigger	0	0	0	0	0	2	1
usher	1	2	0	0	0	0	0
influence	0	0	0	1	1	0	1
interest	0	0	0	0	2	1	0
galvanize	0	0	0	1	0	1	1
condition	0	0	0	0	1	1	0
propel	0	0	0	0	1	1	0
reason	0	0	1	0	0	0	0
impel	0	0	1	0	1	0	0
program	0	0	0	0	0	1	0
motivate	0	0	0	0	0	1	0
induce	0	0	0	0	0	0	1

The normalized frequencies of this group of verbs are presented in Figure 3. As we can see, while there might have been a small increase over time (note the slightly increasing trendline), "neutral" verbs were certainly already present 150–170 years ago, at nearly the same frequency as today.
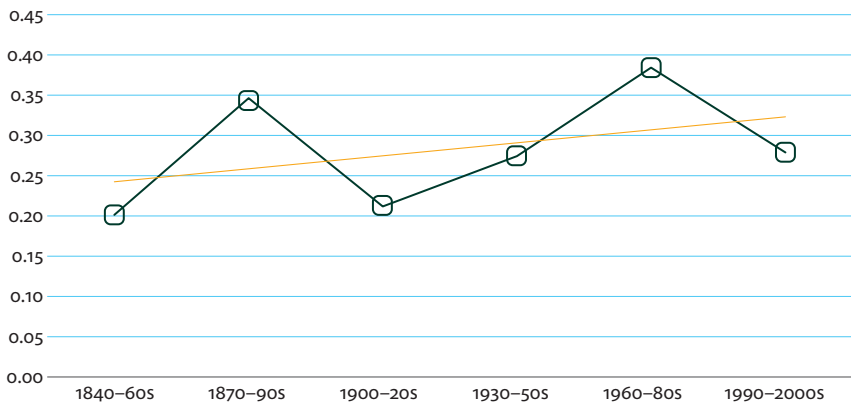
Figure 3. Normalized frequencies of "neutral" verbs in *COHA*

The following are just a handful of tokens with these verbs from the 1840s–1880s. Note that those examples that are marked with an asterisk (*lead* 1815, *usher* 1824 and perhaps *lead* 1835) occur with a V-*ing* form that is more nominal in meaning, but which is potentially ambiguous.

(6) a.   the bold, contending, daring spirit of the world which *leads* mankind into
          quarreling and fighting   (*COHA*, NF, 1815*)
     b.   our habit of disputation in the defence of any side – *leads* us very much
          into doubting   (*COHA*, FIC, 1835*)
     c.   It was not *ushered into being* by the warmth of popular excitement.
                                                                       (*COHA*, MAG, 1824*)
     d.   Heart of mine alone refuseth To be chided, To be *guided into hating* where
          it perished   (*COHA*, FIC, 1848*)
     e.   I was *drawn into speaking* of my life at home.   (*COHA*, FIC, 1862)
     f.   land … can be *stimulated into producing* a meagre green crop of some kind
                                                                       (*COHA*, NF, 1880)
     g.   I have never yet been able to *reason* myself *into feeling* old.
                                                                       (*COHA*, FIC, 1884)

The bottom line, however, is that we can already find many examples of the "*into-causative*" construction with these neutral verbs (especially with *lead*) in the 1800s. Hence, they do not represent a semantic extension of the construction in the past few decades.

7.   Positive verbs

There is one semantic class of verbs where there does appear to be a recent change (and which has not been discussed in previous research), and this is an increase with "positive" verbs like *charm, enchant*, and *romance* (where there is influence from X on Y, but in a positive sense). Table 6 shows the frequency of the "*into*-causative" construction with these verbs in the different time periods, and Figure 4 provides the normalized frequencies (per million words).

Table 6. Frequency distribution of "positive" verbs in *COHA*, by thirty-year period

	1810–30s	1840–60s	1870–90s	1900–20s	1930–50s	1960–80s	1990–2000s
charm	0	0	0	1	6	3	7
jolly	0	0	0	2	2	0	0
enchant	0	0	0	2	0	0	1
urge	0	1	0	0	1	1	0
ease	0	0	0	0	0	1	1
motivate	0	0	0	0	0	1	0
romance	0	0	0	0	0	0	1
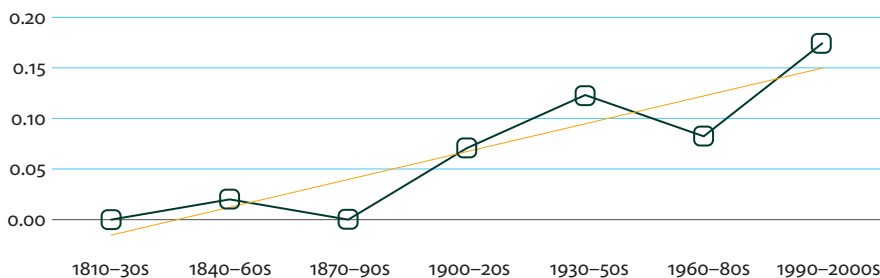
Figure 4. Normalized frequency of "positive" verbs in *COHA*

The following are a handful of examples with these "positive" verbs, and most of them come from more recent decades.

(7)   a.   By talking fast you could sometimes *jolly* the railroads *into cutting* rates way done.   (*COHA*, MAG, 1934)
       b.   you can *ease* the child *into relating* his homework period to the regular activities of others   (*COHA*, NF, 1970)
       c.   or weak parents will be *motivated into doing* a better job of looking after their children   (*COHA*, MAG, 1971)
       d.   That sister of mine could *charm* a marble statue *into carrying* on a pleasant conversation.   (*COHA*, FIC, 1977)

    e.   It's not about *romancing* the client *into buying* the clothes.
                                                       (*COHA*, NEWS, 1995)

    f.   A dreadlocked man with a mellifluous voice *enchants* you *into buying* lavender.
                                                      (*COHA*, NEWS, 2004)

The above examples from *COHA* are supplemented in (8) by examples with other positive verbs from *COCA* (none of which are attested in the 400 million words of *COHA*). (Note that sentences with an asterisk have additional tokens with that verb in *COCA*.)

(8) a.   It wasn't some dusty book that *inspired* Darwin *into thinking* about the order of life in new ways.   (*COCA*, MAG, 1999*)[9]

      b.   Like a mother cat licking her kitten, he *groomed* her back *into trusting* him.
                                              (*COCA*, MAG, 1994*)

      c.   Do you think a gay person could *pray* himself *into being* not gay?
                                              (*COCA*, SPOK, 2002)

      d.   to *love* some of us *into loving* some of them.   (*COCA*, NEWS, 2007*)

The examples of the "*into*-causative" construction with positive verbs are quite interesting, in terms of the semantic extension of the construction. Along with the negative verbs (e.g. *force, mislead, frighten, wheedle*), there have from the beginning also been less negative verbs of persuasion, such as *coax* and the verb *persuade* itself. So it appears that there was always the *possibility* of extending the construction to explicitly positive verbs, such as *motivate, enchant*, and *love*, but it is only in the last few decades that such an extension has actually taken place.

8.   Indirect causation

In semantic terms, the "*into*-causative" construction typically involves direct causation from X on Y, as in:

(9) a.   Bill (X) talked Sue (Y) into paying (Z) for the meal.

      b.   Fred (X) bamboozled his friends (Y) into supporting (Z) his plan.

Within just the past two decades or so, however, the construction has apparently spread semantically to cases where the causation is much more indirect, and this is a topic that has not been discussed in any previous research. In these cases, X

---

9.   The abbreviations of the genre categories in *COCA* are as follows: ACAD = Academic journals, FIC = Fiction, MAG = Popular magazines, NEWS = Newspapers, SPOK = Spoken

merely creates *an environment* in which Y can do something (Z). The following are a handful of examples from *COCA* and *GloWbE-US*:[10]

(10)   a.   It (X) has helped to *build* America (Y) *into exploring* new frontiers (Z).
            (*COCA*, SPOK, 1994)
       b.   My gators start, like seeds, small, but you (X) *feed* them (Y) *into becoming* attractions (Z).   (*COCA*, FIC, 2004)
       c.   Fran (X) also *organized* Kathy (Y) *into buying* (Z) some smart cotton dresses.   (*COCA*, FIC, 1997)
       d.   when his wife *educates* him *into believing* that her staying home is a feminist statement   (*COCA*, MAG, 1998)
       e.   We're trying to *integrate* this person *into being* a law-abiding citizen again.
            (*COCA*, NEWS, 2009)
       f.   so a heading in a list (X) implicitly *splits* the list (Y) *into spanning* multiple sections (Z)   (*GloWbE-US*)

Let us consider some of these examples more closely. In (10a), something (X) has helped America (Y) to be a certain way, and then (indirectly, as a result) America (Y) can explore new frontiers (Z). Similarly in (10b), you (X) feed the gators (Y) and then (indirectly, as a result) they become attractions (Z). There are also cases like (10c), where it is not clear just how much Fran (X) influenced Kathy (Y) to buy some dress (Z). In other words, instances of the "*into*-causative" construction may be positioned on a continuum of force or influence, with the most prototypical sentences like (9) instantiating direct causation at one end of the continuum, and instances of much more indirect causation, as in (10), at the other end.

In these cases of indirect causation, the data for the "*into*-causative" construction fits in very well with Goldberg's (1995) analysis (from a Construction Grammar perspective), in which the *construction* (X verbs Y into Z-*ing*) is the important issue, and essentially any verb could occur as the matrix verb, as long as there is some degree of force or coercion by X on Y.

As far as the historical development of these cases of "indirect causation" is concerned, our corpus evidence suggests that they are a very recent phenomenon. We have individually examined each of the 544 matrix verbs for the "*into*-causative" construction in *COHA*, to identify those that do not belong to one of the prototypical categories such as force, persuasion, coercion, or emotional reaction. Because directness is a gradient phenomenon, it is not possible to say exactly how many

---

10. *GloWbE* is a 1.9 billion word corpus that is based on web pages (many of them blogs) from 20 different English-speaking countries, including 385 million words from the US; see http://corpus2.byu.edu/glowbe for more details.

indirect verbs are involved, but it is probably between 15 to 20 types with a total of about 25 to 30 tokens. Although it is conceivable that there are a handful of examples that we have missed in the 400 million words of text in *COHA*, our research suggests that there are no examples of indirect causation before the 1990s. Since that time, however, they occur quite readily in *COCA* and *GloWbE*, as evidenced by examples like those in (11) above. The construction thus appears to have recently expanded (again, see Goldberg 1995) to include the new pragmatic domain of indirect causation, which was previously not possible.

9.   Interaction with the "*way*" construction

The corpus data also show an interesting interaction between the "*into*-causative" construction and the "*way*" construction, a construction that has been studied by Israel (1996), Goldberg (1995, 1997, 2006), and others. The "*way*" construction is composed of a verb followed by a possessive, the word *way*, and a preposition, as in (11):

(11)   a.   Sally *made her way through* the crowd.
       b.   We *elbowed our way out of* the building.
       c.   It *clawed its way up* the ladder.

As the preceding research indicates, there is nothing about *claw* or *elbow*, for example, that creates a sense in which movement occurs, but rather it is created by the entire construction.

The *way* construction is fairly common in *COHA*, occurring about 18,500 times. Figure 5 shows the normalized frequency of the construction by decade from the 1810s–2000s. As can be seen, the frequencies show a fairly flat development from the early 1800s, with perhaps just a small increase in the last 10 to 20 years.

Figure 5. Frequency of the *way* construction in *COHA* (normalized per million words)

An interesting recent addition to the "*into*-causative" construction is a "hybrid" construction, which is part "*into*-causative" and part *way* construction. Before the 1990s, there was one single example of this hybrid construction in COHA:

(12)   Moscow and Petrograd were off limits except to those few wealthy Jews who could *bribe* their way *into sending* a son or a daughter beyond the boundaries.
(*COHA*, FIC, 1958)

In the last two decades, however, the frequency of the hybrid construction has increased quite a bit. There are 9 tokens in *COCA* and another 17 in *GloWbE-US*. The following are a handful of examples from *COCA*:

(13)   a.   "I had to *imagine* my way *into being* a good mother," she said.
(*COCA*, NEWS, 2009)

b.   and now I have *lied* my way *into having* to leave the house altogether
(*COCA*, FIC, 1996)

c.   what we''re doing is – is *muddling* our way *into making* a war criminal a hero
(*COCA*, SPOK, 1999)

d.   and we would *grow* our way *into destroying* the huge deficits
(*COCA*, SPOK, 1991)

e.   so we need our parents and our moms to *act* their way *into loving* themselves
(*COCA*, SPOK, 2006)

f.   Mr. Zell *talked* his way *into managing* some off-campus housing property.
(*COCA*, NEWS, 2001)

g.   and *talked* his way *into being* admitted to the mission primary school
(*COCA*, 1992, ACAD)

Recall that the semantic structure of the "*into*-causative" construction can be characterized as (14), with (15) as an example:,

(14)   X causes Y to do / become Z

(15)   Bill (X) talked Sue (Y) into coming over (Z)

In that light, the new hybrid construction works well for those contexts in which no specific Y is mentioned, and *way* (*his way, her way, their way*, etc.) is used as a "placeholder" for the missing Y:

(16)   a.   [they] *shouldered* their way *into controlling* the fabulously wealthy gold mines
(*COCA*, MAG, 1999)

b.   Boston's "shortstop of the future" has *belted* his way *into being* a star of the present.
(*COCA*, MAG, 1998)

In (16a), "they" are not "shouldered" into control of the gold mines by anyone else, and in (16b) no one "belted" the shortstop into a condition in which he was considered "a star of the present". And that is the function of the hybrid construction – it works nicely for those contexts in which the other participant (Y) does not (or cannot) exist, or else is not important or relevant.

Perhaps even more unusual are the cases like (17):

(17) a.   "I had to *imagine* my way *into being* a good mother," she said.
                                                        (*COCA*, NEWS, 2009)
      b.   and now I have *lied* my way *into having* to leave the house altogether
                                                        (*COCA*, FIC, 1996)
      c.   he had *stammered* his way *into asking* Lois to dance with him
                                                        (*GloWbE-US*)
      d.   You don't *think* your way *into becoming* yourself.   (*GloWbE-US*)
      e.   *Nose* your way *into doing* extracurricular activity you are interested in.
                                                        (*GloWbE-US*)
      f.   Every civilization …has always been able to *reason* its way *into … denying*
           the most minimal claims to justice.   (*GloWbE-US*)

In these cases, there is no possible Y (separate from X) that could take the place of the *way* phrase. For example, in (17a) it would be difficult to say "I (X) had to imagine Sue (Y) into being a good mother" or in (17c) "He (X) had stammered his brother (Y) into asking Lois to dance with him".

This hybrid construction constitutes an interesting development, as it was essentially non-existent throughout the history of the "*into*-causative" construction until very recently, and then began to appear in just a very narrow semantic/pragmatic context, where the participant Y is not important. One might wonder what other constructions (besides the *way* construction) the "*into*-causative" construction will join forces with in coming decades, as it carves out new semantic niches.

10. Conclusion, and a final note on corpus size

This corpus based study has provided several insights into recent historical shifts with the "*into*-causative" construction. The verbs that occur with this construction over time are probably much more diverse than anyone might have imagined, but the most frequent semantic categories (force, persuasion, deception, etc.) have been the same since at least the early 1800s. There has not been an increase in "neutral" verbs (as some previous research had suggested), but it is only quite recently that the construction has been used with "positive" verbs, and it is likewise only very recently that it has extended to "indirect" causation. Finally, there has been an

interesting interplay between the "*into*-causative construction" and the "*way* construction" over the past few decades.

A recurring theme throughout this paper has been the importance of robust corpus-based data. Using *COHA*, *TIME*, and *COCA*, we have access to more than 5,500 tokens with more than 500 different matrix verbs for the 1810s–1980s, and then another 5,800 or so tokens for the 1990–2000s. Smaller 2–4 million word corpora (such as *ARCHER* or the *Brown* family of corpora) would not have been nearly big enough to provide useful data. They would typically only yield 1/100th of the number of tokens that we have in *COHA*, *TIME*, and *COCA*, and in most cases that is far too small to say anything insightful about the construction. In summary, robust corpora – like those that have only recently become available – allow us to map out a wide range of semantic and syntactic shifts with constructions like the "*into*-causative" in ways that would be quite impossible otherwise.

References

Bridgeman, Lorraine, Dillinger, Dale, Higgins, Constance, Seaman, P. David & Shank, Floyd A. 1965. *More Classes of Verbs in English*. Bloomington IN: Indiana University Linguistics Club.

Davies, Mark. 2009. The 385+ million word *Corpus of Contemporary American English* (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics* 14: 159–190. doi:10.1075/ijcl.14.2.02dav

Davies, Mark. 2010. The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing* 25: 447–464. doi:10.1093/llc/fqq018

Davies, Mark. 2012a. Expanding horizons in historical linguistics with the 400 million word Corpus of Historical American English. *Corpora* 7: 121–157. doi:10.3366/cor.2012.0024

Davies, Mark. 2012b. Some methodological issues related to corpus-based investigations of recent syntactic changes in English. In *The Oxford Handbook of the History of English*, Terttu Nevalainen & Elizabeth Closs Traugott (eds), 157–174. Oxford: OUP.

Francis, Gill, Hunston, Susan & Manning, Elizabeth (eds). 1996. *Collins Cobuild Grammar Patterns, 1: Verbs*. London: HarperCollins.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago IL: University of Chicago Press.

Goldberg, Adele. 1997. Making one's way through the data. In *Grammatical Constructions: Their Form and Meaning*, Masayoshi Shibatani & Sandra A. Thompson (eds), 29–53. Oxford: Clarendon Press.

Goldberg, Adele. 2006. *Constructions at Work: Constructionist Approaches in Context*. Oxford: OUP.

Gries, Stefan T. & Stefanowitsch, Anatol. 2003. Co-varying collexemes in the into-causative. In *Language, Culture, and Mind*, Michel Achard & Suzanne Kemmer (eds), 225–236. Stanford CA: CSLI.

Hunston, Susan & Francis, Gill. 2000. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English* [Studies in Corpus Linguistics 4]. Amsterdam: John Benjamins. doi:10.1075/scl.4

Israel, Michael. 1996. The way constructions grow. In *Conceptual Structure, Discourse and Language*, Adele Goldberg (ed.), 217–230. Stanford CA: CSLI.

Kim, Jong-Bok & Davies, Mark. 2016. The INTO-CAUSATIVE construction in English: A construction-based perspective. *English Language and Linguistics* 20: 55–83

Mair, Christian. 2002. Three changing patterns of verb complementation in Late Modern English: A real-time study based on matching text corpora. *English Language and Linguistics* 6: 106–131. doi:10.1017/S1360674302001065

Rohdenburg, Günter. 2007. Functional constraints in syntactic change: The rise and fall of prepositional constructions in Early and Late Modern English. *English Studies* 88: 217–233 doi:10.1080/00138380601042824

Rudanko, Juhani. 1991. On verbs governing in *-ing* in present-day English. *English Studies* 72: 55–72. doi:10.1080/00138389108598733

Rudanko, Juhani. 2000. *Corpora and Complementation: Tracing Sentential Complementation Patterns of Nouns, Adjectives, and Verbs over the Last Three Centuries*. Lanham MD: University Press of America.

Rudanko, Juhani. 2002. *Complements and Constructions*. Lanham MD: University Press of America.

Rudanko, Juhani. 2005. Lexico-grammatical innovation in current British and American English: A case study on the transitive *into -ing* pattern with evidence from the Bank of English Corpus. *Studia Neophilologica* 77: 171–187. doi:10.1080/00393270500384791

Rudanko, Juhani. 2006. Emergent alternation in complement selection: The spread of the transitive 'into – ing' construction in British and American English. *Journal of English Linguistics* 34: 312–331. doi:10.1177/0075424206293600

Rudanko, Juhani. 2011. *Changes in Complementation in British and American English*. Houndmills: Palgrave Macmillan. doi:10.1057/9780230305199

Rudanko, Juhani. 2015a. 'Wheedled me into lending him my best hunter': Comparing the emergence of the transitive into -ing construction in British and American English. In *Perspectives on Complementation: Structure, Variation and Boundaries*, Mikko Höglund, Paul Rickman, Juhani Rudanko & Jukka Havu (eds), 128–140. London: Palgrave Macmillan.

Rudanko, Juhani. 2015b. *Linking Form and Meaning: Studies on Selected Control Patterns in Recent English*. London: Palgrave Macmillan.

Vosberg, Uwe. 2003. Cognitive complexity and the establishment of *-ing* constructions with retrospective verbs in Modern English. In *Insights into Late Modern English*, Marina Dossena & Charles Jones (eds), 197–220. Bern: Peter Lang.

Wulff, Stephanie, Stefanowitsch, Anatol & Gries, Stefan. 2007. Brutal Brits and persuasive Americans: Variety-specific meaning construction in the *into*-causative. In *Aspects of Meaning Construction*, Gunther Radden, Klaus-Michaël Köpcke, Thomas Berg & Peter Siemund (eds), 265–281. Amsterdam: John Benjamins. doi:10.1075/z.136.17wul

Free adjuncts in Late Modern English

A corpus-based study

Carla Bouzada-Jabois

University of Vigo

The present paper focuses on the English free adjunct construction in Late Modern English based on data from the *Penn Parsed Corpus of Modern British English* (*PPCMBE1*). Data from this analysis will be compared to information available in existing studies on Early Modern and Present-day English. The findings based on my own data analysis and the data in earlier studies confirm a decrease in the frequency of the free adjunct construction from EModE to LModE and an stabilization from LModE to PDE. Verbal free adjuncts show an increasing preference for present-participial forms and they tend to be placed in final position. Semantically, most verbal free adjuncts are controlled by the subjects of the main clauses and their semantic contribution to the sentence is, in Kortmann's (1991: 121) terms, "most informative". Finally, free adjuncts are typically not augmented.

Keywords: free adjuncts, adverbial, dependency, control

1.   Introduction

This paper addresses the development of free adjuncts (FA) from Early Modern English (EModE) to Present-day English (PDE), with a focus on changes in the Late Modern English (LModE) period. FAs, as in (1) below, are typically nonfinite constructions conveying adverbial meaning with respect to a main clause.

(1)   They were going away, and we were going after them, *firing at them too*.
                                    (*PPCMBE1*, HOLMES-TRIAL-1749,69.1289)

FAs are sometimes labeled extra-clausal constituents or supplements (Dik 1997: 381; Huddleston & Pullum 2002: 1265) because they are detached from the main clause, usually by means of punctuation (or intonation in speech). In other words, they are not fully integrated into the clause structure, and in general they are not essential

constituents in the syntactic structure of the clause. They may also take up different positions in the clause. Semantically, they establish a referential link with an element or constituent in the main clause, often the main clause subject, and their meaning is usually indeterminate and needs to be inferred from the context.

FAs may have a verbal or a non-verbal head (Kortmann 1991: 6). Non-verbal FAs, however, are not as common as verbal FAs (Kortmann 1995: 195); in fact, it is largely the verbal types that are in focus in the relevant literature (see, e.g., Río-Rey 2002; Killie & Swan 2009). Accordingly, non-verbal FAs, as in (2), will not be included in this study:

(2)   *A virtuoso in the art of the discourteous aside*, he had never been subjected to
       such disrespect.   (example taken from Kortmann 1991: 6)

FAs have often been examined in relation to absolute constructions, in that both constructions share a number of structural, distributional, and semantic properties. What distinguishes them is the expression of the subject: while FAs lack explicit subjects, these are compulsory in absolutes. There is general agreement in the literature that the implicit subject of an FA by default equals the main clause subject, while the explicit subject of an absolute differs from the main clause subject (Visser 1972: 1132, 1147; Kortmann 1991: 5; Río-Rey 2002: 311); see example (3), a prototypical example of a verbal absolute construction.

(3)   I slept very ill, and read W. M. in the quiet hours between 3.0 and 5.0, *the weir
       rushing outside*, the night still and cold. (*PPCMBE1*, BENSON-190X,131.796)

The remainder of this paper is structured as follows. In Section 2, I provide a theoretical description of FAs, looking at types of heads (2.1), the position FAs occupy within the matrix clause (2.2), the elements acting as introducers (2.3), semantic features such as subject control and adverbial meaning (2.4), and textual distribution (2.5). Section 3 includes a description of the corpus and summarizes the methodology. Section 4, then, discusses the findings for each of the characteristics introduced in Section 2. Conclusions are offered in Section 5.

2.   An introduction to FAs and their characteristics

In the following sections, I present the features in terms of which FAs have been described in the literature and which also underlie the data analysis in Section 4. Special attention will be paid here to the description of these features in Kortmann (1991).

2.1   Head of the free adjunct

As I already pointed out, this paper focuses on verbal FAs. They can be further classified according to the form of their non-finite head, which may be a present participle (4), past participle (5), or infinitive (6) (Kortmann 1991: 6). Participial predicates may also be marked for perfect aspect (7), passive voice (5), or a combination of the two (8), as well as for progressive aspect (9) (see Kortmann 1991: 7; Kortmann 1995: 193). In the present study, the data have been categorized following the classification in Kortmann (1995: 195) for verbal FAs; that means that FAs with perfect passive participles, as in (8), and with progressive passive participles, as in (9), have been classified together with other past participial FAs marked for passive voice, as in (5). Furthermore, infinitival FAs have not been considered, mainly for reasons of frequency: they only take up 1.4 percent of the FAs in Kortmann (1995: 195). Examples (4) to (7) below were taken from Kortmann (1991: 6–7) while (8) to (9) are taken from the *PPCMBE1*.

(4)   *Inflating her lungs*, Mary screamed.

(5)   *Compared with other European countries*, the German inflation rate is very low.

(6)   *To finance this adventures*, Americans are digging deeper into their pockets.

(7)   *Having failed the exam twice*, he didn't want to try again.

(8)   About Two we came to Trewint, wet and weary enough, *having been battered by the Rain and Hail for some Hours.*   (*PPCMBE1*, WESLEY-174X,21.231)

(9)   but *being dispers'd by a Storm*, it was the 19th of July before they enter'd the Channel;   (*PPCMBE1*, KIMBER-1742,269.C2.270)

2.2   Position

FAs can occupy different positions with respect to the main clause. Following common practice (Quirk et al. 1985: 490; Huddleston & Pullum 2002: 779–784; Hasselgård 2010: 42), this study distinguishes three main positions: initial position, as in (10), when FAs precede the main clause; final position, as in (11), when FAs appear after the main clause; and medial position, where FAs interrupt the main clause at some point. Although FAs in medial position most often occur "immediately after their antecedent"[1] (Quirk et al. 1985: 1125; Declerck 1991: 457), as in (12), other instances of medial position are attested in my data: an example is (13), where the FA occurs in between the auxiliary and the main verb. Beukema (1982: 71)

---

1.   "Antecedent" is Quirk et al.'s (1985: 1125) label for Kortmann's (1991, 1995) "controller", in this case the subject of the main clause.

and Stump (1985: 6) restrict medial position to a location after a non-pronominal subject, as in (12), yet examples like (14) show that medial position may also occur following a pronominal subject.

(10) *Having its ultimate roots in the human mind, like a great many other sciences*, it is a recondite branch of the vast subject of Sociology, or Society, viewed both as structure and as history. (*PPCMBE1*, BAIN-1878,378.285)

(11) it was taken in the fields, in the beginning of April, *running amongst the leaves and dead grass*. (*PPCMBE1*, ALBIN-1736,3.51)

(12) Fitzpatrick, *hanging down his Head*, repeated, 'that he had committed a Mistake, for which he heartily asked Pardon,' (*PPCMBE1*, FIELDING-1749,3,10.399)

(13) every Freeholder, …, before he is admitted to poll at the said Election, shall (*if required by the Candidates, or any of them, or any other Persons having a Right to vote at the said Election*) first take the Oath … following; (*PPCMBE1*, STATUTES-1745,6,266.153)

(14) These, *when corrected*, not only serve as the best versions to be translated back into Latin, but bring children to a greater proficiency in their own language. (*PPCMBE1*, BARCLAY-1743,99.325)

The position of FAs with respect to the main clause is likely to reflect their discourse-pragmatic function (Kortmann 1995: 228). Adverbial clauses in initial position are said to "serve a frame-setting function for the material that follows" (Kortmann 1995: 228), while those occupying final positions "serve a much more 'local' function by exhibiting a high degree of referential continuity with the matrix clause and by providing more specific or additional information to the matrix proposition" (Kortmann 1995: 228–229). These two positions are described in Chafe (1984: 448) respectively as "guideposts" orienting the reader or as "afterthoughts" commenting on the information in the preceding main clause.

Word order in FAs is also claimed to be iconic (Stump 1985: 321; Kortmann 1995: 117). FAs preceding the main clause tend to refer to events occurring before the event in the main clause, while FAs following the main clause are expected to denote events subsequent to the event in the main clause. Temporal succession may also receive a non-temporal interpretation (Kortmann 1991: 116). Events occurring before the event in the main clause are usually understood as the cause/reason or the condition for the main event to take place. For FAs occurring in end position, a non-temporal interpretation is likely to indicate a consequence, result or purpose of the situation in the main clause. The interpretation of FAs will be examined in Section 2.4 below and some comments on the relation between their position and meaning will be made in Section 4.5.

2.3   Elements introducing free adjuncts

FAs can be introduced by elements which further specify the semantic relation with the main clause, namely subordinating conjunctions or prepositions. These introductory elements have been labeled "augmentors" (Stump 1985). Examples of so-called "augmented" FAs introduced by a conjunction are (13)–(14) above and (15)–(16) below; FAs introduced by a preposition are (17)–(18). Note that conjunctions introducing an FA may take *-ing* as well as *-ed* complements, while prepositions only take *-ing* complements.

(15)   for, *though directed to Downing-street*, it would not, as other letters would have done, address itself to the present possessor.

(*PPCMBE1*, WALPOLE-174X,5,18.439)

(16)   But, like a brave Soldier in the first Action wherein he is engaged, he continued resolute, *tho' shuddering at the Terror of the Assault*;

(*PPCMBE1*, DODDRIDGE-1747,47.356)

(17)   *Before returning home*, she bought presents for her parents.

(example taken from Kortmann 1991: 8)

(18)   Besides, *in talking with their young companions*, they learn to speak in a distinct manner.   (*PPCMBE1*, BARCLAY-1743,21.99)

Whereas augmentation of FAs by subordinating conjunctions is widely acknowledged in the literature (Kruisinga 1932: 274–280; Visser 1972: 1133–1139, 1255–1257; Bäcklund 1984; Stump 1985; Declerck 1991: 457; Kortmann 1991: 7–8), augmentation by prepositions of *-ing* FAs is less accepted. Kortmann (1991: 8, 196) accepts FAs introduced by *before* and *after* due to their ambivalent status as prepositions as well as conjunctions (see also Bäcklund 1984: 51 as regards *since*, *until*, and *till*), but he contends (1991: 181) that while examples of "nominal *-ing* clauses of the '*in/on/upon/by* + V-*ing*' type (…) display similar properties as free adjuncts", they "are nevertheless to be kept separate", thus rejecting examples such as (18) above.

Linguists accepting prepositions as introductory elements of FAs include Kruisinga (1932: 277), Visser (1972: 1133–1139), Declerck (1991: 36, 457), and Pérez-Quintero (2002: 36); they usually provide examples such as (19) and (20):

(19)   *On stepping out of the car* she felt a fine rain in her face.

(example taken from Visser 1972: 1136)

(20)   Are you sure you're not making a big mistake *in selling these shares*?

(example taken from Declerck 1991: 458)

The reason why linguists are hesitant to accept *-ing* forms introduced by prepositions as FAs can be attributed to the traditional distinction whereby conjunctions

combine with (non-finite) verb forms and prepositions with nominal forms. This point of view is exemplified in Kortmann (1995: 199–201): while he admits that "adverbial participles" and "adverbial gerunds" have "parallel semantic and pragmatic behaviour" (1995: 201) – indeed, they are both adverbials – he states that the *-ing* form itself in examples such as (20) assumes a nominal function and should therefore be considered different from an FA.[2]

It can, however, be argued that *on hearing the news* in (21) is also an FA, because it is interchangeable with *hearing the news*, an FA by all traditional criteria. Similarly, Declerck (1991: 457) paraphrases the adverbial *on being told* in (22) as *When she was told*…, thus underscoring that an FA with *-ing* introduced by a preposition functions in the same way as an FA with a finite verb form.

(21)   *On hearing the news*, she immediately telephoned her father.
                                (example taken from Huddleston & Pullum 2002: 1262)

(22)   She was very relieved *on being told that she was no longer needed*.
                                (example taken from Declerck 1991: 458)

In this study, FAs augmented by both conjunctions and prepositions are examined on the basis of their functional similarities. In treating FAs introduced by conjunctions on a par with FAs introduced by a preposition, this study aligns itself with Huddleston & Pullum's (2002: 1220–1222) view, according to which "there is no difference of form, function, or interpretation that correlates systematically with the traditional distinction between 'gerund' and 'present participle'". The gerund/participle distinction will not be further pursued in the present study (for a more in depth analysis of this issue, see De Smet 2010, 2014).

2.4   Semantic features

The semantic characterization of FAs encompassed in this section addresses (i) the co-referential links that hold between the FA and the main clause and (ii) the classification of FAs according to their semantic contribution to the sentence.

It has been claimed in the literature that the implicit subject in an FA, as a rule, corresponds to the explicit main clause subject (Visser 1972: 1132; Declerck 1991: 36, 456; Kortmann 1991: 5; Río-Rey 2002: 311). FAs such as (23) are thus categorized as

---

2.   Even though the *-ing* form in examples such as (17), repeated here as (i), can also fulfill a nominal function, Kortmann (1995: 200) acknowledges that *before* as augmentor can introduce not only nominal elements but also finite clauses. In the latter case it functions as a subordinating conjunction and not as a preposition.

(i)   *Before returning home*, she bought presents for her parents.   (Kortmann 1991: 8)

"related" (Kortmann 1991: 43). When the subject of the FA does not correspond to the subject of the main clause, as in examples (24) and (25), the FA is regarded as "unrelated" (Kortmann 1991: 43).

(23) Then, *crossing the Thames*, he pacified the Eastern Counties,
(*PPCMBE1*, OMAN-1895,400.467)

(24) and then, *being unfit for Company*, I sent *him* home Tongue-tied,
(*PPCMBE1*, STEVENS-1745,23.169)

(25) *Looking westward across the narrows and over the wide, turf-clad, gorse-sprinkled, sandy flats*, … the steep and rugged promontory of Penmaenbach shoots finely up eight hundred feet into the sky,   (*PPCMBE1*, BRADLEY-1905,199.27)

With respect to their semantic contribution to the superordinate clause, FAs have been described as "having the force of full adverbial clauses" (Curme 1947: 150) backgrounding some information in relation to the main clause (Berent 1973: 147; Thompson 1983: 44; Stump 1985: 1). However, the meaning FAs convey is often not very specific, but rather underdetermined (Stump 1985: 1; Kortmann 1991: 1; Haspelmath 1995: 27), and their connection with the main clause needs to be inferred from the context.

The semantic categorization of FAs is not always straightforward: in the absence of augmentors specifying or restricting the range of possible interpretations of an FA (Kruisinga 1932: 279; Stump 1985: 13, 322; Kortmann 1991: 194), its interpretation needs to be determined on an individual basis. The interpretation attributed to a given FA depends on factors such as the semantics of the main clause to which it is attached, the particular context in which the utterance takes place, or the knowledge of the speaker and/or hearer involved in the communication process (Stump 1985: 321). An example of an FA with variable interpretation is given in (26) below, where either a causality or a concession interpretation may be involved. When several meanings can be associated with an FA, we will follow Kortmann (1991: 133, 1995: 215) in ascribing the construction to the most informative type (see below).

(26) *Knowing the importance of his evidence*, the witness preferred to stay at home.
(example taken from Kortmann 1991: 2)

In Kortmann (1991: 121), a typology of 15 semantic categories is set up. These semantic categories range from most to least informative depending on the degree of "interpretational complexity" (Fonteyn & van de Pol 2016: 10), that is, on the amount of knowledge, "or (co- or contextually substantiated) evidence" (Kortmann 1991: 119–120) that is required to be identified as the semantic relation holding between the FA and the clause. Kortmann (1991: 120) argues that the least informative meanings simply add some information to the main clause, while the most informative relations

modify the meaning of the main clause. Least informative interpretations include addition, accompanying circumstance, exemplification, specification, same time, and manner; the most informative ones include anteriority and posteriority, cause, result, purpose, condition, contrast, and concession. The following examples illustrate relations of accompanying circumstance (27), manner (28), anteriority (29), cause (30), conditionality (31), and concession (32).[3]

(27)   She then goes out, *followed by the servant, who closes the door after him.*
(*PPCMBE1*, WILDE-1895,69.816)

(28)   At the Island where the Ships were, they found fresh Water, *digging two or three Foot down into the Sand.*   (*PPCMBE1*, COOKE-1712,1,435.253)

(29)   Xerxes march'd on, laying all waste before him; and *being arriv'd at Panopaea, a Town in Phocis*, he divided his Army in two parts,
(*PPCMBE1*, HIND-1707,316.207)

(30)   *hearing nothing of them*, the next Morning, the Boat went off again,
(*PPCMBE1*, COOKE-1712,1,441.338)

(31)   'You shall swear (or *being one of the People called Quakers*, you shall solemnly affirm) that you are a Freeholder in the County of and have a Freehold Estate, consisting of (…)   (*PPCMBE1*, STATUTES-1745,6,266.154)

(32)   He asked us whether he should try and make for the nearest of the Azores, Corvo, and run the ship on shore, or whether he should make for the harbour of Fayal, *incurring the risk of sinking in doing so.*
(*PPCMBE1*, FAYRER-1900,18.479)

2.5   Textual distribution

FAs have been claimed to be more frequent in written than in spoken language (Thompson 1983: 45; Kortmann 1991: 2; Río-Rey 2002: 313) and more common in less formal text types of a more narrative nature (Kortmann 1991: 2). Thompson (1983: 46) points out that FAs abound in "discourse that attempts to describe by creating an image", and Kortmann in fact (1995: 191) confirms that "it is the *depictive* versus *nondepictive* distinction which is relevant" to the study of FAs in written registers.

---

3.   The semantic interpretation given to these examples is context-dependent and, as explained above, if more than one interpretation is available, the most informative one is selected (Kortmann 1991: 133, 1995: 215).

3.   Data and methodology

This study of FAs is corpus-based and examples have been retrieved from the first edition of the *Penn Parsed Corpus of Modern British English* (*PPCMBE1*) (Kroch et al. 2010), a compilation of texts comprising around one million words covering the period 1700–1914. The *PPCMBE1* is divided into three seventy-year periods: 1700–1769 (298,764 words), 1770–1839 (368,804 words), and 1840–1914 (281,327 words). Only the first and last period have been selected for analysis as a way of recognizing more easily any developments within LModE. These two periods will be referred to as P1 and P3, respectively. The corpus contains 18 different text types, from which Bible excerpts have been discarded in order to avoid skewing of the data due to archaic and formulaic syntax.

The *PPCMBE1* is syntactically annotated. A precision and recall analysis on a sample of the corpus shows that 99.4 percent of the examples of FAs were automatically retrieved by using the parsing conventions, and even though a process of extensive manual pruning was necessary (only 38.8 percent of the examples retrieved automatically correspond to FAs), the corpus tagging proves to be reliable.

To widen the scope of the diachronic study, I have also considered two quantitative studies based on Early Modern English (EModE) and PDE data. Río-Rey (2002) analyzes the period 1500 to 1710 and a total of seven genres from the EModE section of the *Helsinki Corpus*. Comparison with the PDE data has been carried out by taking into account Kortmann's (1991) seminal monograph, in which he compiled his own corpus comprising the period 1975–1987, and covering four text types: fiction, news, science, and spoken material. Since the above studies do not employ exactly the same features as I do here, some restrictions will be needed for comparison across time, and these will be acknowledged when necessary.

4.   Analysis of the data and results

In this section, I present the results of my analysis of FA in LModE, and I compare my results with data from the other empirical studies of FAs in EModE (Rio-Rey 2002) and PDE (Kortmann 1991). My discussion is structured in terms of the characteristics presented in Section 2.

4.1 Overall frequencies

A total of 1,935 examples were retrieved from the two subperiods of the *PPCMBE1*, 1,076 for P1, and 859 for P3. Figure 1 plots the normalized frequencies (frequency per 10,000 words) in LModE as well as the frequencies provided for EModE in Río-Rey (2002) and for PDE in Kortmann (1991); Río-Rey's (2002: 314) database contains 847 FAs in EModE, and Kortmann (1995: 1326) analyzes 1,326 verbal FAs.[4]
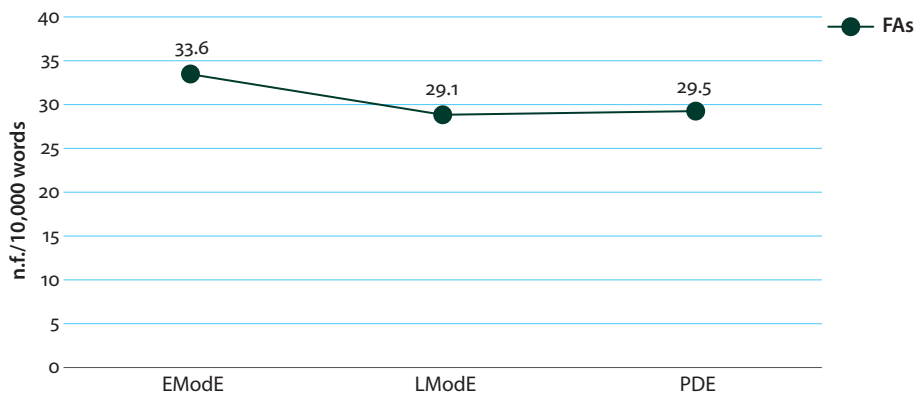
Figure 1. Diachronic distribution of FAs from EModE to PDE[5]

The decrease of FAs from EModE to PDE is significant overall ($\chi^2(2) = 12.5$, $P = 0.0019$). Whereas the decrease from EModE to LModE proves to be statistically very significant ($\chi^2(1) = 11.11$, $P = 0.0009$), Figure 1 illustrates a period of stabilization in the evolution of FAs from LModE to PDE. The evolution from LModE to PDE is not statistically significant ($\chi^2(1) = 0.09$, $P = 0.7642$).

---

4. The total number of FAs in Kortmann (1991, 1995) is 1,412, but this figure also includes non-verbal types. It has been possible to omit non-verbal and infinitival FAs from the total number of FAs since Kortmann (1995: 195) provides specific percentages for each of the FA-types. It has not been possible, however, to exclude the non-verbal or the infinitival examples in the subsequent discussion of the individual variables. Since non-verbal and infinitival FAs account only for 6.1 percent (86 examples) in Kortmann's (1991, 1995) database, the statistical impact of their inclusion in my discussion of the ensuing variables will not alter the overall results of my analysis.

5. In order to make an accurate comparison with Kortmann (1991, 1995) and Río-Rey's (2002) results, FAs augmented by prepositions have been excluded from consideration in this figure (see Sections 2.3 and 4.4 for more information).

4.2 Head elements

In a number of earlier studies on FAs in PDE (Kruisinga 1932: 276; Kortmann 1991: 120, 1995: 195), the *-ing* form (present participle) is claimed to be the most frequent type of verbal head. The relative share of the different types of verbal head investigated in my LModE data are shown in Figure 2.
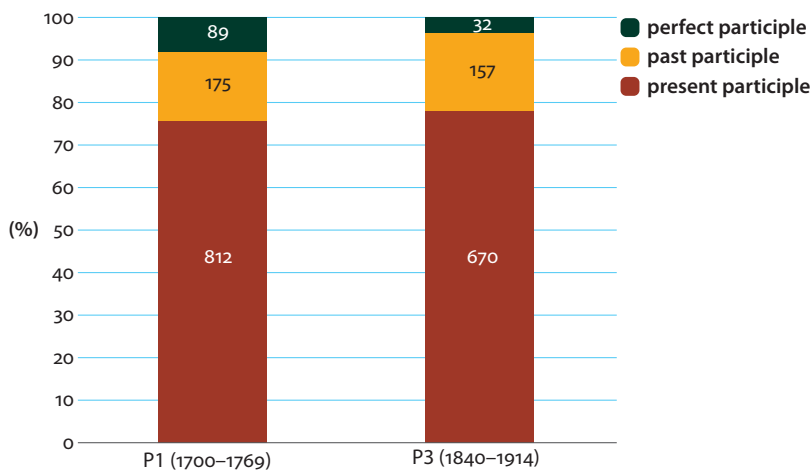
Figure 2. Verbal head elements in LModE

With 75.5% and 78% of the FAs in P1 and P3, respectively, my data corroborate these earlier claims on the preference for present participial heads. Past participle FAs represent about 16% and 18% of all verbal FAs in P1 and P3, followed by perfect participle forms, which account for only around 8% and 4%.

Figure 3, then, presents the changing distribution of verbal head types from LModE to PDE (based on Kortmann 1995). This development does not include data from EModE since Río-Rey (2002) does not provide information about verbal heads in her study; the development shown can therefore only be partial.

Figure 3 shows that 76.6% of the FAs have a present-participle head in LModE and that this share increases significantly to 95.7% in PDE ($\chi^2(2) = 237.49$, $P < .0001$). At the same time, perfect participles and, most notably, past participles decrease drastically over time – if the raw figures were normalized per 10,000 words, the proportion of FAs headed by a past/perfect participle would not even amount to one attestation per 10,000 words in PDE.

This noticeable increase of present participles has also been pointed out in studies on absolute constructions. Van de Pol & Cuyckens (2013: 220), for instance, report that from EModE to PDE the present participle type is by far the most frequent head in absolute constructions.
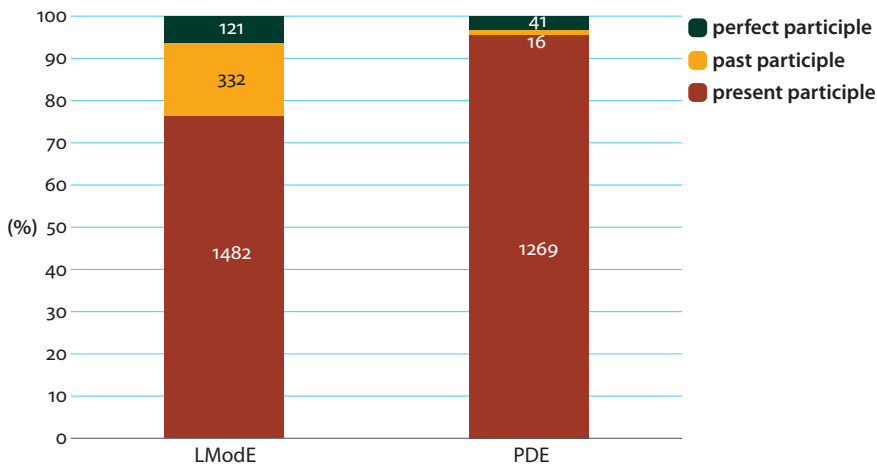
Figure 3. Distribution of FAs depending on their verbal head

4.3 Position

As pointed out in Section 2.2, there are three different positions available for FAs in the sentence: initial, medial, and final. The positional distribution of FAs in LModE and PDE (data taken from Kortmann 1991: 139) is given in Figure 4, no information being available concerning position in EModE. The two LModE subperiods are represented separately to allow any diachronic developments to emerge.
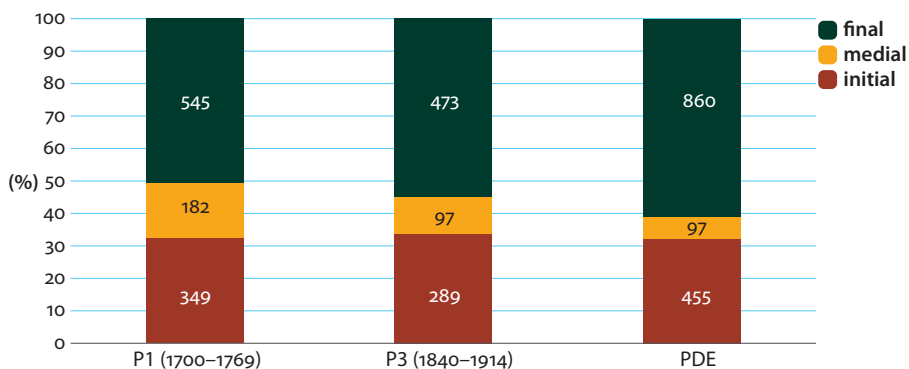
Figure 4. Position of FAs in LModE and PDE

Figure 4 shows that there is an increasing tendency towards FAs in sentence-final position, from 50.7% in P1 (the first period in LModE) to 60.9% in PDE; this increase is statistically significant ($\chi^2(2) = 66.88$, $P < .0001$). Since the proportion

of sentence-initial FAs across the various periods remains largely constant, the increase of FAs in final position correlates with the decrease of medial FAs, that is, of those interrupting the clause (from 16.9% to 11.3% and 6.9% in each of the respective periods).

A pilot comparison of the evolution of FAs and absolute constructions in LModE and PDE, carried out in Bouzada-Jabois & Pérez-Guerra (2014), also evinced a notable preference for sentence-final absolute constructions (from 50 percent in P1 to more than 80 percent in PDE, an increase in line with the results reported in Kortmann 1991: 139). These results corroborate previous claims on the decrease of the relative frequency of sentences with initial adverbials or adjuncts since the Old English period (Breivik & Swan 1994: 28; Pérez-Guerra 1999: 220–221).

4.4   Introductory elements

Kortmann (1991: 195) claims that the presence of introductory elements (or augmentors) in FAs is not a prototypical feature of the construction. Figure 5 sets out my findings on augmentation for LModE. The augmentors here include both conjunctions and prepositions even though, as pointed out in Section 2.3, the inclusion of prepositions as augmentors is controversial.
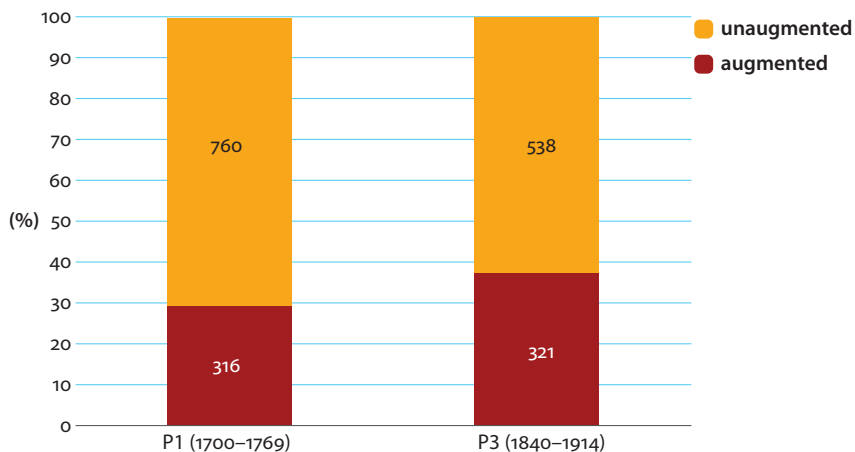
Figure 5. Augmentation of FAs in LModE

Figure 5 shows that augmentation is the preferred option in 29.4% of the FAs in P1, and increases to 37.4% in P3; this increase is statistically significant ($\chi^2(1) = 13.49$, $P = 0.0002$). The results after excluding the controversial prepositionally augmented FAs are shown in Figure 6.
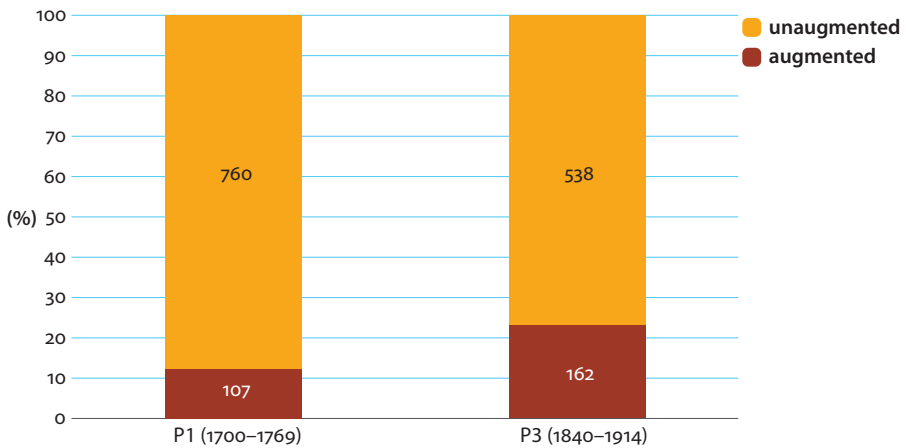
Figure 6. Augmentation of FAs in LModE (prepositional elements excluded)

The results in Figure 6 confirm a proportional increase in augmented FAs also when conjunctions are considered as the only means of augmentation. This increase from 12.3% in P1 to 23.1% in P3 is statistically significant ($\chi^2(1) = 31.02$, $P < .0001$). Figure 7 provides information on the two groups of augmentors. The first group includes conjunctions and those ambivalent augmentors mentioned in Kortmann (1991, 1995) that may function as prepositions and the second group includes prepositional elements.
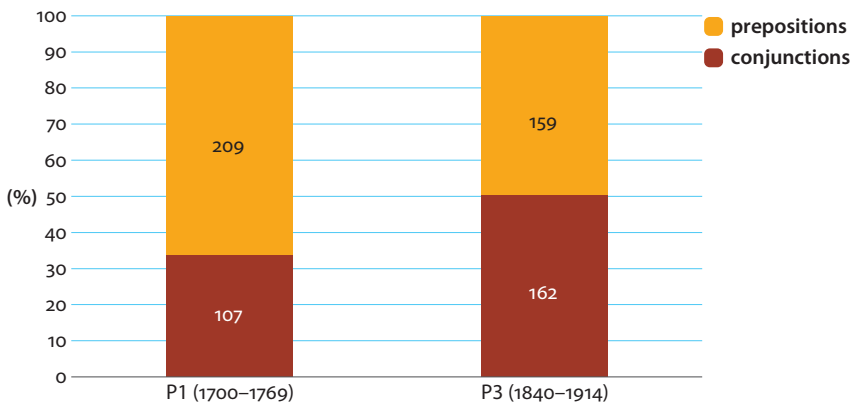
Figure 7. Augmentation types in LModE

On the one hand, the rate of prepositionally augmented FAs accounts for 66.1% of the examples in P1 and decreases to 49.6% in P3. The reason for this high proportion of FAs with introductory prepositions may be the high token frequency

of the preposition *by*, which accounts for 45.9% of all the augmented examples in P1 and 55.3% in P3. On the other hand, the proportion of FAs introduced by conjunctions increases from 33.9% in P1 to 50.5% in P3; this is a significant increase ($\chi^2(1) = 17.33$, $P < .0001$). In order to examine whether augmentation by conjunctions further increases into PDE, Kortmann's frequency data on augmented FAs (which do not include prepositionally augmented FAs) are compared with my LModE data, with the prepositionally augmented FAs removed. The relevant results for this comparison are given in Figure 8.
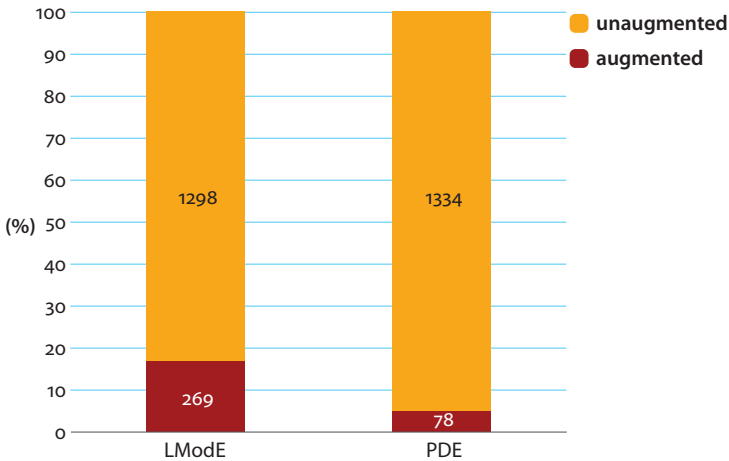
Figure 8. Augmentation in FAs (prepositional elements excluded) in LModE and PDE

Contrary to expectations, the proportion of FAs augmented by conjunctions is shown to decrease significantly ($\chi^2(1) = 96.7$, $P < .0001$) between LModE and PDE, from 17.2% in LModE to 5.5% in PDE. This may either show that the increase in augmented examples between P1 and P3, illustrated in Figure 5, did not continue into PDE, or that in PDE augmentation is more productive with prepositional introducers.[6]

---

6. In their investigation of FAs in the *ICE-GB* corpus, Pérez-Guerra & Bouzada-Jabois (2015) considered every example introduced by conjunctive and prepositional augmentors. The data evince a tendency towards augmentation from LModE to PDE when the prepositional examples are included.

4.5   Semantics

In this section, which addresses the semantics of FAs, I will first deal with subject control links between the FAs and the main clauses, and then describe a number of findings on the semantic classification of FAs.

As I explained in Section 2.4, FAs which maintain a referential link with the subject of the main clause have been coded as "related FAs". By contrast, those FAs which hold a referential link with another constituent or which have no referent in the main clause have been labeled as "unrelated FAs". Figure 9 shows the proportions of related and unrelated FAs from EModE to PDE.
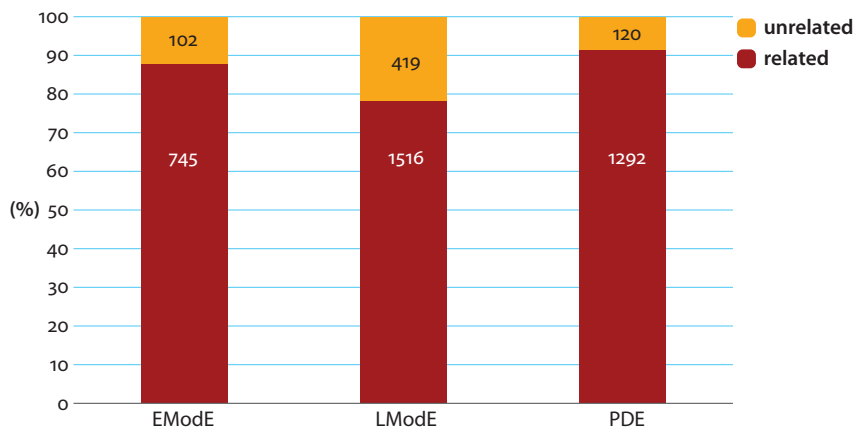
Figure 9. Subject control in FAs

The results in Figure 9 show that, across the various periods under consideration, most FAs are of the related type (i.e. their unexpressed subjects are controlled by the subject of the main clause). Unrelated FAs account for 12% in EModE, 21.7% in LModE, and 8.5% in PDE (see Kortmann 1991: 48). When comparing the frequencies in EModE and PDE, the evolution can be said to be statistically significant ($\chi^2(1) = 7.11$, $P = 0.0077$).

When considering relatedness against position, unrelated FAs occur most frequently in initial position. When the FA precedes the subject, it may have a connective function to the previous context rather than to the main clause subject following the FA (with which it would be expected to establish co-reference). An example is shown in (33), where the subject of the FA should be looked for in the previous context.

(33)   [*Wheat* cut about a fortnight before it is ripe contains the most starch and gluten; the bushel weighs heavier, and the straw contains its greatest nourishment;] *cut late*, the ear contains more cellulose, consequently an increased production of bran, and a diminished proportion of flour.   (*PPCMBE1*, FLEMING-1886,71.56)

An FA in final position is far removed from the main clause subject constituent and this does not facilitate coreference between the FA and the subject, yet the subject has already been mentioned in the discourse and this makes its selection as a referent more plausible than when the FA is in initial position. Finally, related FAs occurring in medial position favor the related type to the greatest degree. Medial FAs tend to appear immediately after the subject of the main clause, bringing about coreference between the medial FA and the subject.

As regards the semantic typology of FAs, I have adopted Kortmann's (1991: 121–133) classification in an attempt to facilitate the comparison of the data between LModE and PDE. Since Río-Rey (2002) does not provide information on the interpretation of FAs, the focus here will be on LModE and PDE, as in Figure 10.
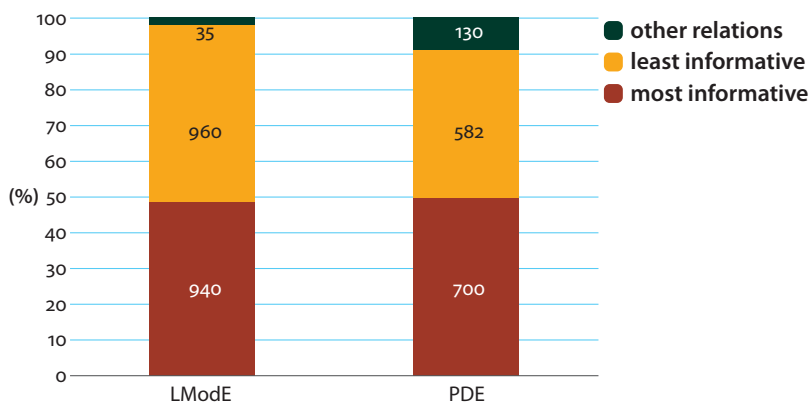
Figure 10. Diachronic evolution of the meaning of FAs

Figure 10 shows that, on the one hand, the relative frequency of the most informative types, such as (34) below conveying a concessive meaning, is very similar in both periods (48.6% and 49.6% in LModE and PDE, respectively). On the other hand, the least informative meanings, such as (35) below indicating simultaneity, decrease from LModE to PDE as a result of an increase of the so-called "other relations".[7] While the overall evolution of the semantics of the construction is statistically significant ($\chi^2(2) = 103.28$, $P < .0001$), the most informative FAs do not undergo any statistically significant change ($\chi^2(1) = 0.29$, $P = 0.5902$) over time.

---

7. The category "other relations" includes the following: instances in which it is not possible to select only one meaning, examples which have been grouped by Kortmann (1991: 133) in a specific category that is not included in his taxonomy since they require obligatory augmentation (FAs of comparison, substitution, etc.), and cases which are problematic for other reasons (see Kortmann 1991: 133–134 in this respect).

(34)   Cortes … *tho' disswaded by the Tlascalans*, concluded to take his Way through the City of Cholula …   (*PPCMBE1*, COOKE-1712,1,422.66)

(35)   Four weeks before admission, *whilst engaged at his work on board ship*, had a sudden attack of acute pain, starting from the cardiac region and then coursing down the left arm as far as the fingers.   (*PPCMBE1*, POORE-1876,166.72)

Previous studies (Kortmann 1995: 228; Killie & Swan 2009; Fonteyn & van de Pol 2016: 11) have claimed that FAs preferentially express more informative relations. My findings show that in LModE the relative frequencies of the more and of the less informative semantic relations are very similar. It is only in PDE that the FAs conveying more informative relations become the preferred semantic type.

As I pointed out in Section 2.2, there is a correlation between the semantic relation expressed by an FA and its position with respect to the main clause. Figure 11 shows the proportions of the three different semantic interpretations in relation to the position they occupy with respect to their main clause.
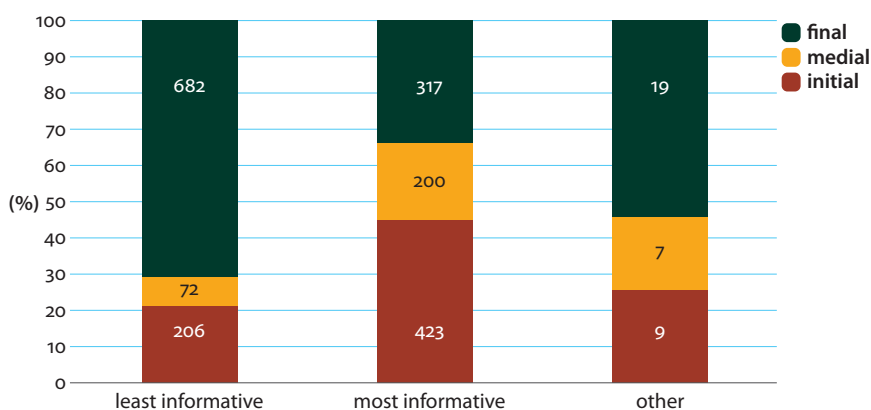
Figure 11. The relation between semantic relation and the position of FAs in LModE

On the one hand, the least informative relations favor end position FAs, which account for 71% of the FAs with less informative meanings. In contrast, the most informative relations favor initial position FAs; these account for 45% of all the FAs conveying most informative relations. There are few examples of "other relations" in LModE, yet these favor final position.

The position that the FA occupies with respect to the main clause can be said to attest to word-order iconicity, i.e. the order of words reflects the succession of events (Stump 1985: 321; Kortmann 1991: 117). The most frequent FAs in initial position are those conveying cause, condition, and anteriority (44%, 41%, and 62%, respectively), all of them referring to events preceding the action in the main clause. For end position, the most common semantic relations are accompanying circumstance,
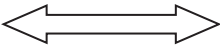
manner, and simultaneity (87%, 66%, and 55%, respectively). A clear example of the iconicity of word order in final position is illustrated by FAs with a resultative interpretation, such as (36). The resultative examples appear all in final position and word-order reflects clearly the succession of events.

(36)   In jumping a burn I slipped and fell with violence on my left arm, *hurting it severely.* (*PPCMBE1*, FAYRER-1900,4.73)

4.6   Textual distribution

As I mentioned in Section 2.5, FAs are more frequent in written than in spoken registers and show a preference for less formal, narrative registers. The 17 text types from the *PPCMBE1* analyzed for this paper have been classified here into two broad categories, adapted from Culpeper & Kÿto (2010: 18), namely "writing-related" and "speech-related" types. A third category "fiction" has been added as an in-between text type. This text type comprises both dialogue and narrative material, which does not belong in a clear-cut fashion to either of the two major textual categories (indeed, adding the FAs retrieved from fictional texts to either of the major textual (writing-/speech-related) categories would compromise the quality of the data). Table 1 sketches the classification of the text types in the *PPCMBE1* corpus.

Table 1. Writing- and speech-related text types

Writing-related (formal; to be read)		Speech-related (speech-like/purposed/based)
biography_auto, biography_other, education_treatise, handbook_other, history, law, philosophy, science_medicine, science_other, travelogue	fiction	diary, drama_comedy, letter_non-private, letter_private, proceedings_trial, sermon

The distribution of FAs in LModE according to this classification of text types is provided in Figure 12.

In both subperiods, FAs occur much more frequently in writing-related than in speech-related registers, thus confirming the claim about the high productivity of FAs in written registers (Thompson 1983: 45; Kortmann 1991: 2; Río-Rey 2002: 313). This can even be observed in P3, where FAs are also becoming more frequent in fictional text types, with a corresponding decrease in speech-related registers. The changing distribution of FAs in the different text-types is statistically significant ($\chi^2(2) = 22.3$, $P < .0001$) from P1 to P3.
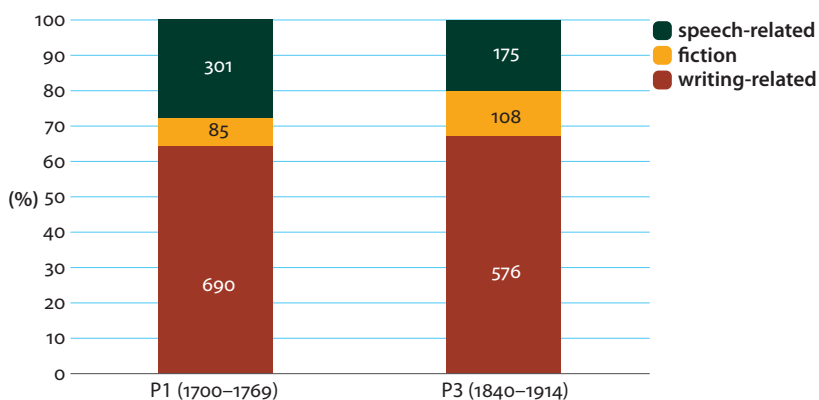
Figure 12.   Frequency of FAs in writing- versus speech-related text types in LModE

Comparisons of the distribution of FAs across text types between my LModE data and Early Modern and PDE are complicated by a disparity in the text types distinguished in the three data sets or quantitative studies considered in this paper. Río-Rey's (2002: 314) seven text types (fiction, travelogue, science, private letters, law, sermons, and comedies) in EModE might be easily transferable to the three categories in my study; Kortmann's (1991) selection for PDE, however, is not broad enough for such an adaptation, in that he only provides data from fiction, news, science, and spoken material. Therefore, Figure 13 plots only my results for LModE and Río-Rey's (2002) for EModE.
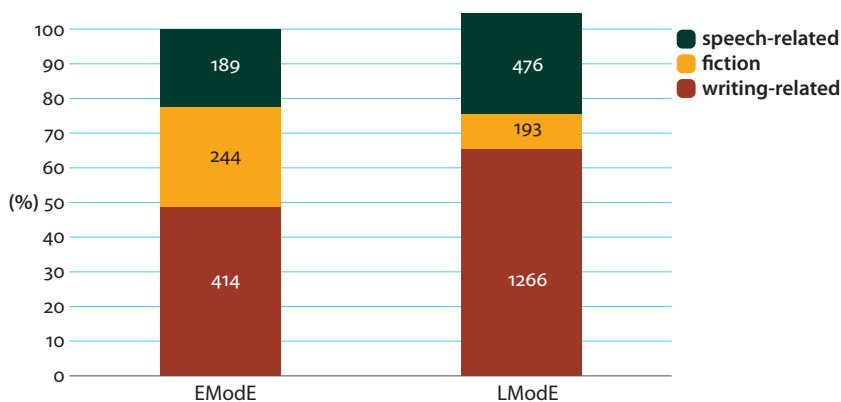
Figure 13.   Diachronic evolution of FAs per text type

Even though the results can only be tentative, for reasons already discussed, Figure 13 shows that FAs are more frequent in writing-related text types than in the spoken material. FAs in both written and spoken registers grow significantly from EModE

to LModE ($\chi^2(2)$ = 161.03, P < .0001), whereas those in the fiction genre decrease in LModE.

5.   Conclusions

The following conclusions can be drawn on the basis of our own analysis of LModE data, as well as of data in other empirical studies (Rio-Rey 2002 for EModE; Kortmann 1991, 1995 for PDE). First, the data have shown a decrease of the FA construction from EModE to LModE and a stabilization from LModE to PDE. Second, regarding head elements, present-participial FAs are by far the most frequent type, accounting for 76.6% of the FAs in my data for LModE and reaching 95.7% of the verbal FAs in PDE (Kortmann 1991, 1995). Third, a clear tendency towards placement in sentence-final position has been observed from LModE to PDE. Fourth, introductory elements (augmentors) are not prototypical in LModE and PDE – information about augmentors in EModE was not available in the empirical study we employed. Fifth, as expected, most of the FAs are semantically related to the subject constituent in the main clause. Moreover, the number of related FAs shows an increase towards PDE. Sixth, with respect to the semantic content of FAs, the data have not revealed significant differences from LModE onwards concerning the relative frequency of most informative semantic relations. Finally, this paper has corroborated previous claims on the productivity of FAs in written-related text types. However, any comparisons with earlier and later periods of the language are tentative, and require a more accurate typology of text types.

Summarizing, this paper has focused on the analysis of the FA construction in LModE paying attention to the elements heading the construction, the position FAs occupy with respect to the main clause, augmentation, and the semantics of the construction, both as regards referential links and adverbial meaning. The discussion has included a comparison of the results obtained for LModE with those of previous studies of FAs in EModE and PDE.

Acknowledgements

References

Bäcklund, Ingegerd. 1984. *Conjunction-headed Abbreviated Clauses in English. Acta Universitatis Upsaliensis* [*Studia Anglistica Upsaliensia* 50]. Uppsala: Almqvist & Wiksell.

Berent, Gerald P. 1973. Absolute constructions as 'subordinate clauses'. In *You Take the High Node and I'll Take the Low Node: Papers from the Comparative Syntax Festival*, Claudia Corum, Cedric Smith-Stark & Ann Weiser (eds), 147–154. Chicago IL: Chicago Linguistic Society.

Beukema, Frits H. 1982. On the internal structure of free adjuncts. In *Linguistics in the Netherlands 1982*, Saskia Daalder & Marinel Gerritsen (eds), 71–82. Amsterdam: North-Holland.

Bouzada-Jabois, Carla & Pérez-Guerra, Javier. 2014. Regularising the nonfinite periphery in English: Absolutes, free adjuncts and constructionalisation in Modern English. Paper delivered at the 38th International Conference of AEDEAN, Alcalá de Henares, 12–14 November.

Breivik, Leiv Egil & Swan, Toril. 1994. Initial adverbials and word order in English with special reference to the Early Modern English Period. In *Studies in Early Modern English*, Dieter Kastovsky (ed.), 11–43. Berlin: Mouton de Gruyter. doi:10.1515/9783110879599.11

Chafe, Wallace. 1984. How people use adverbial clauses. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, Claudia Brugman & Monica Macaulay (eds), 437–449. Berkeley CA: Berkeley Linguistics Society.

Culpeper, Jonathan & Kytö, Merja. 2010. *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: CUP.

Curme, George O. 1947. *English Grammar*. New York NY: Barnes & Noble.

Declerck, Renaat. 1991. *A Comprehensive Descriptive Grammar of English*. Tokyo: Kaitakusha.

De Smet, Hendrik. 2010. English *-ing*-clauses and their problems: The structure of grammatical categories. *Linguistics* 48(6): 1153–1193. doi:10.1515/ling.2010.038

De Smet, Hendrik. 2014. Constrained confusion: The gerund/participle distinction in Late Modern English. In *Late Modern English Syntax*, Marianne Hundt (ed.), 224–238. Cambridge: CUP. doi:10.1017/CBO9781139507226.017

Dik, Simon C. 1997. *The Theory of Functional Grammar, Part 2: Complex and Derived Constructions*. Berlin: Mouton de Gruyter.

Fonteyn, Lauren & van de Pol, Nikki. 2016. Divide and conquer: The formation and functional dynamics of the Modern English *ing*-clause network. *English Language and Linguistics* 20(2): 185–219. doi:10.1017/S1360674315000258

Hasselgård, Hilde. 2010. *Adjunct Adverbials in English*. Cambridge: CUP. doi:10.1017/CBO9780511676253

Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In *Converbs in cross-linguistic perspective: Structure and Meaning of Adverbial Verb Forms – Adverbial Participles and Gerunds*, Martin Haspelmath & Ekkehard König (eds), 1–55. Berlin: Mouton de Gruyter.

Huddleston, Rodney & Pullum, Geoffrey K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.

Killie, Kristin & Swan, Toril. 2009. The grammaticalization and subjectification of adverbial *-ing* (converb clauses) in English. *English Language and Linguistics* 13: 337–363. doi:10.1017/S1360674309990141

Kortmann, Bernd. 1991. *Free Adjuncts and Absolutes in English: Problems of Control and Interpretation*. London: Routledge.

Kortmann, Bernd. 1995. Adverbial participial clauses in English. In *Converbs in cross-linguistic perspective: Structure and Meaning of Adverbial Verb Forms – Adverbial Participles and Gerunds*, Martin Haspelmath & Ekkehard König (eds), 189–237. Berlin: Mouton de Gruyter.

Kroch, Anthony, Santorini, Beatrice & Diertani, Ariel (comp.). 2010. *Penn Parsed Corpus of Modern British English*.

Kruisinga, Etsko. 1932. *A Handbook of Present Day English, Vol. II: English Accidence and Syntax 3*. Groningen: P. Noordhoff.

Pérez-Guerra, Javier. 1999. *Historical English Syntax: A Statistical Corpus-based Study on the Organisation of Early Modern English Sentences*. Munich: Lincom.

Pérez-Guerra, Javier & Bouzada-Jabois, Carla. 2015. Patterning verbal absolutes and free adjuncts in Contemporary English, with a grain of history. Paper delivered at the 6th Biennial International Conference on the Linguistics of Contemporary English (B-ICLCE 6), Madison, WI, 19–23 August.

Pérez Quintero, María Jesús. 2002. *Adverbial Subordination in English: A Functional Approach*. Amsterdam: Rodopi.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Río-Rey, Carmen. 2002. Subject control and coreference in Early Modern English free adjuncts and absolutes. *English Language and Linguistics* 6(2): 309–323. doi:10.1017/S1360674302000254

Stump, Gregory T. 1985. *The Semantic Variability of Absolute Constructions*. Dordrecht: Reidel. doi:10.1007/978-94-009-5277-5

Thompson, Sandra A. 1983. Grammar and discourse: The English detached participial clause. In *Discourse Perspectives on Syntax*, Flora Klein-Andreu (ed.), 43–65. New York NY: Academic Press.

van de Pol, Nikki & Cuyckens, Hubert. 2013. Gradualness in change in English augmented absolutes. In *Synchrony and Diachrony: A Dynamic Interface* [Studies in Language Companion Series 133], Anna Giacalone Ramat, Caterina Mauri & Piera Molinelli (eds), 341–366. Amsterdam: John Benjamins. doi:10.1075/slcs.133.14pol

Visser, Fredericus Theodorus. 1972. *An Historical Syntax of the English Language, Part two: Syntactical Units with One Verb*. Leiden: E.J. Brill.

CHAPTER 8

Complexity and genre distribution of left-dislocated strings after the fixation of SVO syntax

David Tizón-Couto
Universidade de Vigo

This paper investigates the diachronic development of strings that include both a left-dislocated constituent and a coreferring resumptive in the subsequent clause, thus resembling the reportedly speech-like and informal contemporary Left Dislocation construction. The data is analyzed according to a range of factors relating to the inner configuration of such strings and examined by means of the pertinent statistical tests. The purpose is to arrive at a clearer and broader picture of (i) the general decline of left-dislocated strings, (ii) the factors determining their structural complexity, and (iii) their role as prospective markers of orality in Modern English speech-related texts.

Keywords: Left Dislocation, complexity, orality, genre, Modern English

1. Introduction

Left Dislocation (LD) is the term that scholars of generative grammar (and especially of syntactic movement) have used to describe the construction in (1). It was first labeled "left dislocation" by Ross (1967: 253),[1] and it has been more recently described by Netz and Kuzar (2007) as a marked theme construction (such as Object Fronting: *Beans he likes*). In particular, LD involves a syntactically disconnected item (*this girl* in (1)) that is positioned to the left of the main proposition and linked to that main proposition by means of a coreferential resumptive item (see Pérez-Guerra 1999; Traugott 2007; Pérez-Guerra & Tizón-Couto 2009; Tizón-Couto 2012, 2015).

---

1. See Anagnostopoulou et al. (1997) for a collection of key papers on the syntactic analysis of LD.

© 2018 John Benjamins Publishing Company

(1)   *This girl* this morning **she** threw a wobbly.   (Biber et al. 1999: 956)[2]

English written historical texts contain a variety of LD-strategies, as can be seen in (2)–(7):

(2)   Diccon: *Hodge*, **he** intendes this same night, to slip in there awayes.
      (*Penn-Helskinki Parsed Corpus of Early Modern English* (*PPCEME*), drama, E1)[3]

(3)   *Your sea-scratchers of Holland in the Indies*, neither ar we muche behoulded to
      them at home, though they have their beeing by us;
                    (*Parsed Corpus of Early English Correspondence* (*PCEEC*), letter, E2)

(4)   *Mrs. Man, and her and good Mr. Panninan*, **thy all** in queare very much of you,
                    (*PCEEC*, letter, E3)

(5)   *She that was Maryed*, **she** had vpon her hede a crowne of gold.
                    (*PPCEME*, travelogue, E1)

(6)   Throckmorton: *Whatsoeuer Wiat hath saide of me in hope of his Life*, he vnsayde
      it at his Death.   (*PPCEME*, trial, E1)

(7)   So *Tom seeing them bait at him in such a manner as they did*, **he** went first to
      one work, then to another,   (*PPCEME*, fiction, E3)

The left-dislocated (LDed) slot is held by a one-word NP in (2), an NP postmodi-fied by a PP in (3), a series of NPs separated by commas in (4), an NP modified by a relative clause in (5), a *wh*-clause in (6), and an *ing*-clause with an overt subject that corefers with the following *he* in (7). The LDed item in (6) has been defined as a "Free Relative clause" (Andrews 2007: 213–214), and in (7), it illustrates what Kortmann (1991, 1995) terms the "Absolute Construction" (which can – but need not – occur in sentence-initial position; see Río-Rey 2002; Fonteyn & van de Pol 2016). As opening constituents that corefer with a resumptive in the main clause, both (6) and (7) meet the required characteristics for LD. Although these examples are included in the current analysis because they bear similarities to dislocation, their status as borderline cases must be made explicit here: the italicized structure in (6) could also be analyzed as a concessive adverbial clause, in analogy to examples like *Whatever I did, I couldn't please him*. Similarly, the italicized structure in (7) can be analyzed as an adverbial clause.

---

2.   In numbered examples, italics are employed to signal the left-detached constituent, while bold indicates the resumptive item.

3.   Citation of tokens includes three labels: (a) name of the corpus, (b) genre and (c) period: E1 (1500–1569), E2 (1570–1639), E3 (1640–1710), L1 (1700–1769), L2 (1770–1839), L3 (1840–1914).

This paper looks into the development of the LD construction in the Modern English (ModE) period (1500–1914), that is, after the fixation of sentence constituents in the SVO pattern (Fischer 1992: 371; van Hoorick 1994: 53). The general aim is to offer a picture of the chronological development in different text genres from ME to LModE. Specifically, the paper will first be concerned with the factors determining the structural complexity of the LDed string; to that effect, a linear regression model will be fitted. Second, it will be investigated to what extent LDed NPs constitute markers of orality or informality. This will be analyzed by means of univariate statistical tests.

An aspect of LD that is frequently neglected is the role of syntactic complexity: Traugott (2007: 435) pointed out the lack of a "discussion of complexity in the literature on left-dislocations" and suggested that "[t]he relation of left-dislocation to the length of the initial nominal [i.e. LDed nominal] and to reference tracking is worthy of study, as is its relation to written or literate vs. spoken or vernacular register". Previous diachronic studies covering periods before 1914 report that LD typically features rather long LDed constituents (a mean of 7–9 words) (Traugott 2007: 415–416; Pérez-Guerra & Tizón-Couto 2009: 42; Tizón-Couto 2015: 45). In contrast, Snider (2005: 23) reports that in contemporary spoken English, "[grammatical] weight distribution peaks at 3 words and then decreases". An approximate mean of 5 to 5.5 words can be inferred from Snider's results on LD in the SWITCHBOARD corpus of contemporary telephone conversations.[4]

In connection with such previously reported results on length, one of the assumptions behind the present study is that longer LDed items will be harder to process syntactically and, thus, that some accommodation strategies may be observed taking the shape of predictor variables that determine the length of the LDed constituent (see Section 3.1). For instance, it will be investigated whether intervening material is unlikely between a long LDed constituent and the ensuing main clause (or, in other words, whether intervening material predicts a shorter LDed constituent). A second hypothesis to be investigated is that, as in contemporary speech, shorter LDed items are expected to be more frequent in historical genres closer to, or resembling, the oral variant (see Section 3.2).

In sum, by zooming in on the role of complexity, this paper explores questions pertaining to (i) the factors that predict the length of the LDed item (under the

---

4. Snider (2005: 18) does not report a mean value for length (or "weight" in his terminology) of the LDed strings in the SWITCHBOARD corpus; however, he includes a histogram showing that 25% of the tokens are three words long and that, roughly, 65% of the tokens are five words long or shorter. An approximate mean has been calculated by taking Snider's histogram into account that ranges between 5 and 5.5 words.

assumption that a lengthier LDed item will be harder to parse) (cf. Hawkins 1992; Rohdenburg 1995, 1996) and (ii) the association between (short and/or unaugmented) LDed constituents and speech-related genres (cf. Culpeper & Kytö 2010).

This paper has been divided into six sections. Section 2 offers a brief summary of previous research on the diachronic evolution of LD. Section 3 spells out the two main goals and hypotheses of the study. Section 4 describes the data, factors, and statistical methods employed. Section 5 presents the findings that relate to each hypothesis and, last, Section 6 provides a brief discussion followed by the concluding remarks.

2.   A brief historical survey of LD

Until the fifteenth century, English was a particular type of verb-second [V2] language (Los 2009: 97). V2 syntax is typically achieved by means of two syntactic operations:

> The finite verb moves into second position, and a second rule topicalizes a constituent from the clause into first position. This constituent may be moved from any position in the clause, may have any syntactic function (subject, object, adjunct), and may have any information-structural status (given, new, unmarked topic, marked focus, etc.). (Los & Komen 2012: 884)

At the time when Old English syntax displayed a topic-first rather than a subject-first arrangement, there was no correspondence between the syntactic function and the information status of the constituent in preverbal position. In particular, this preverbal position was multifunctional in that it could host unmarked topics as well as focused material (Los 2009: 104; Los & Komen 2012; Pérez-Guerra 2012). With regard to LD constructions, then, OE appears to have had two kinds, as in the present-day West-Germanic languages (Anagnostopoulou 1997; Van Riemsdijk & Zwarts 1997): (i) a contrastive type in which the resumptive occupies the preverbal multifunctional position in the clause (CLD) and (ii) a hanging type where the resumptive occupies its base position in the clause and is not contrastive (HTLD; see (8)). In fact, both options are still available: although contemporary English is sometimes quoted as exclusively featuring HTLD (cf. Shaer et al. 2009: 7), CLD is still instantiated when the resumptive both occupies the preverbal subject (or, marginally, object) position and is contrastively stressed, as in (9) (cf. Los and Komen 2012: 894; van Kemenade & Los 2013: 220). Example (10), which will certainly seem odd (or choppy) to contemporary readers, illustrates one of the 13 tokens in the Modern English dataset investigated where an object demonstrative takes the first position (after the LDed constituent) for contrast.

(8) B: Both my husband and I work, and our children are sixth, fourth, and third grade. And the school years are wonderful, they're just wonderful.

A: Uh-huh.

B: *The kids*, **they** are real people and they are interesting and,

(Gregory & Michaelis 2001: 1683)

(9) *The people who earn millions and pay next to no tax*, **those/they** are our targets.

(Birner & Ward 2002: 1413)

(10) Now since all Languages are naturally equal to us, therefore *the first Language we hear*, **that** we shall first understand; (*PPCMBE*, educational treatise, L1)

The loss of V2, by the end of the ME period (around 1450 for both Traugott 1972 and van Kemenade 1987), meant the loss of the first (or preverbal) multifunctional position that could host both unmarked topics and focused material. This accounts for the decrease of constructions that marked contrast in the preverbal position, such as CLD (or focus-marker *only*; cf. Los 2009; Los & Komen 2012: 891). Also, the loss of V2 reduced the options for positioning non-subject constituents and, thus, conditioned information structure. In turn, this led to the rise of new constructions (such as clefts) and a redefinition of older positions (Los & Komen 2012: 896). This redefinition translated into non-canonical orders acquiring specific information-structural functions. Pérez-Guerra (2012: 122), for instance, works on the "assumption that the (marked) patterns which did not conform to the canonical subject-verb-object word order (e.g. Topicalization, LD, and *there*-sentences) survived as escape hatches" or special constructions to deal with objects that are not new (Topicalization) or subjects that are not given (LD and *there*-sentences). Thus, LDs in OE "serve to keep track of referents that are introduced as salient in some respect in the flow of discourse", they "may be discourse-new or discourse-old", "they typically do not introduce a referent for topic continuity", and they are therefore "considerably less constrained than Present Day English left-dislocations are said to be" as regards information structure (Traugott 2007: 436). However, due to the fixation of the SVO pattern and the disappearance of topic-first alternatives, HTLD became a marked (i.e. low frequency) information-structural strategy devoted to easing the processing of an irrecoverable referent in the discourse and/or highlighting it as a new topic (Prince 1997: 123; Gregory & Michaelis 2001; Birner & Ward 2002).[5]

---

5. The multifunctionality of LD has been demonstrated by Geluykens (1992, 1993), Prince (1997, 1998), and Manetta (2007), among others. The construction may be used for marking the introduction of new elements to simplify discourse processing (i), marking particular members of a set that is already active in previous discourse for reference (ii), or marking unexpected subjects, i.e. a subject that does not match the subject of the previous sentence (underlined in (iii)).

As a result of the establishment of the SVO pattern, both CLD and HTLD are reported to have undergone a steep diachronic decline from OE to Present-day English (Los 2009: 114; Pérez-Guerra & Tizón-Couto 2009: 38; Los & Komen 2012: 896), but for different reasons. Firstly, the decrease of CLD must be connected to the general decline in the marking of contrast in the preverbal position; it only survived when the expression of contrast respects SVO, i.e. when the resumptive occupies the subject-first position. The decline of CLD might also be associated to "the development of written conventions" away from "paratactic constructions" (van Kemenade & Los 2013: 221). Secondly, the decrease of HTLD goes hand in hand with its new status as a marked thematic alternative that mostly fronts irrecoverable referents. In spite of its general status as a highly marked strategy that features a left-detached constituent, LD is reported to predominantly respect the SVO pattern in the main clause, as in *My dad, he loves London* (see Section 3.1).

3. Goals and hypotheses

In addition to the description of the chronological development of the LD construction, this paper has two specific goals: to look into the factors that determine the complexity of the LDed constituent (as reflected by its length), and to explore the extent to which shorter and/or unaugmented LDed constituents correlate with speech-related genres. The following subsections provide details as regards these two specific goals and the corresponding hypotheses.

(i) My sister got stabbed. She died. Two of my sisters were living together on 18th street. They had gone to bed, and this man, their girlfriend's husband, came in. He started fussing with my sister and she started to scream. *The landlady*, **she** went up, and he laid her out. So sister went to get a wash cloth to put on her...   (Prince 1997: 122)

(ii) There are many groups of cacti worthy of collection. Even [*opuntias, the plants which tend to give cacti a bad name, with their nasty little barbed hairs or glochids, which are used for 'itching powder', and sharp, barbed spines which go into flesh much more easily than they come* out]$_i$, even **they**$_i$ have much to offer and can make an interesting – if forbidding – collection.   (Prince 1997: 124)

(iii) He didn't need the money... He said, I want you to buy it 'cause I know you'll keep it open. <u>My dad</u> talked to the guy. And *the guy who owned it*, **he** got a loan.
          (Manetta 2007: 1032)

Beyond these specific information-structural roles, LD has also often been claimed to fulfill an overarching topic-marking function supported by the fact that (a) it most frequently fronts elements that are resumed by grammatical subjects and (b) it promotes referents that persist prominently in the ensuing stretch of discourse (cf. Gregory & Michaelis 2001; Birner & Ward 2002; Traugott 2007; Tizón-Couto 2012, 2015: 40, 2016).

3.1   Factors that determine the complexity of the LDed constituent

In studying the factors that may have an effect on the length of the LDed constituent, I will first consider structural properties. To the extent that structural complexity features (e.g. length, intervening material, punctuation, augmentation) correlate with processing complexity, the issue addressed here also amounts to which factors will give rise to a higher, or lower, degree of processing complexity of the LD string (as a whole). Additional textual properties that will be considered here as potential predictors of the form/length of the LDed constituent are periodization and genre cluster.

Previous investigations on the syntactic properties of LD report that resumptive items typically occupy the subject position within the main clause both in spoken English (cf. Gregory & Michaelis 2001: 9; Snider 2005: 17) and in written historical texts (Traugott 2007: 415–416; Tizón-Couto 2015: 42). However, the resumptive slot may also be held by an object, a prepositional complement, or an adverbial (Pérez-Guerra 1999: 199). Because of positional proximity (Hawkins 2004: 37) and the low-cost processing of a pronominal resumptive (Gibson 1998), it could be hypothesized that subject-first and pronominal resumptives (instead of a full NP, for instance) will allow coreference with a more complex LDed constituent.

In ModE, LDed strings might include intervening material between the detached constituent and the following clause (see Section 4.1.2 for examples).[6] Assuming that constructions are shaped by conflicting forces of economy and explicitness (Diessel 2005: 449), two markedly long recognition domains (a long LDed NP and a long intervening constituent) should delay the parsing of the referential link between the fronted referent and the resumptive within the main clause. According to this line of thinking, a negative correlation can be expected between the length of the LDed constituent and its distance to the ensuing clause containing the resumptive element.

Instances of "unaugmented" LDed NPs, i.e. not preceded by a previous grammatical word (see (1) in Section 1), are typically cited in studies that highlight the informal and speech-like character of LD (see Section 3.2). However, the effect of a preceding item on LD is not typically examined. Previous research has indicated that "conjunctionless (or unaugmented) adverbial clauses have a much longer recognition domain and are harder to parse than augmented ones, in which the parser immediately encounters the mother node in the form of a linking conjunction or preposition" (Fonteyn & van de Pol 2016: 197; see also Diessel 2005). As pointed out by these authors, this distinction is expected to be more pronounced when the

---

6. The variable of intervening material has featured in a different domain of English grammar, i.e. complementation. Specifically, Rohdenburg (1995, 1996) showed that the number of words intervening between a matrix predicate and its clausal complement partly determines the formal type of complement (*that*-clause or *to*-infinitive clause).

adverbial clause occupies the first position. This difference in processing complexity of left-detached items leads to the hypothesis that augmentation, i.e. featuring a preceding conjunction, preposition, or adverb, might facilitate the parsing of the referential link between the resumptive and a longer LDed constituent.

3.2   The link between less complex LDed items and speech-related genres

A second issue to be investigated is the extent to which shorter and unaugmented LDed constituents are characteristic of speech-related genres. It has been pointed out that, in Present-day English, LDed NPs are only expected in unplanned discourse or "very loose and informal speech" (Quirk et al. 1985: 1310), they "are almost entirely absent in the formal planned register" (Givón 1979: 229), and may seem "inappropriate in formal registers" (Lambrecht 1994: 182). In fact, contemporary English conversation appears to show the greatest use of LD, while written English shows usage of the device mostly within pseudo conversations (Geluykens 1992: 99; Ono & Thompson 1994). Similarly, Biber et al. (1999: 957) suggest that these "prefaces" (i.e. LDed NPs) are "almost exclusively conversational features". More recently, Leonarduzzi and Herry (2005: 9) have pointed out that LDed NPs have been perceived as markers of informal, substandard, or non-native varieties by a range of both literature and script writers. Against the background of these claims, this paper investigates the link between speech-related genres and some of the features characterizing LD in studies that focus on the spoken variant, such as the prevalence of the unaugmented form of the LDed NP (cf. Ono & Thompson 1994; Hidalgo 2000; Fox et al. 2003) and its shorter length (cf. Snider 2005).

4.   Data and methodology

The data analyzed in this study were extracted automatically, using *CorpusSearch 2*, from the *Penn-Helsinki Parsed corpora of Middle English* (*PPCME2*), *Early Modern English* (*PPCEME*), *Modern British English* (*PPCMBE*), and the *Parsed Corpus of Early English Correspondence* (*PCEEC*).[7] The tokens extracted are headed by an LDed item, i.e. the "LFD" tag in the parsed corpora, which is coreferential with a "resumptive" constituent that carries out a syntactic function in the ensuing clause,

---

7.   To the author's knowledge, only the *Penn-Helsinki Parsed suite* offers machine-searchable tags for Left Dislocation and Resumptive. A combined search for two tags, namely "LFD" and "RSP", when dominated by either "IP" or "CP", permitted the extraction of all the items. Coreferentiality between the two tags is marked by numeric subscripts in the parsed files of these corpora.

i.e. the "RSP" tag in the parsed corpora. The corpus search for the "LFD" tag was limited to the "LFD-NP" subtag because LDed NPs head the construction that has often been cited as characteristic of spoken discourse (see Section 3.2). Note that not all of the examples coded as "LFD-NP" in the corpora contain a typical NP: some are in fact headed by a Free Relative Clause or an Absolute Construction (see examples (6) and (7) in Section 1). However, the original tagging in these tokens has been generally respected in the analysis due to their formal and (possibly) functional commonalities with the LD construction. Still, a few instances of Topicalization and of the "possessive dative" construction (*My neece St John her little one hath bine very ill*, *PCEEC*, E2; cf. Dons 2004: 45) were removed from the dataset.

In order to provide a comprehensive chronological picture of the LD-phenomenon, normalized frequencies and genre distributions are presented for every period in the corpora (see Figure 1, Section 5.1). However, since ModE is the main focus of this paper, a detailed analysis of the variables listed in the following Subsections (5.2 and 5.3) is only applied to the data from EModE and LModE (but not to the data from the *PPCME2* and the two brief ME sections of the *PCEEC*).

Two different analyses are carried out on the Modern English data. First, a linear regression model is fitted to test the continuous variable "length" (i.e. length of the LDed constituent) for the effects of the seven predictor variables introduced below (see Section 4.1). Second, univariate statistical tests (Pearson's chi-squared tests and Kendall's Correlation tests)[8] are employed in order to test the categorical variable "genre cluster" for the effects of the four predictor variables specified in Section 4.2.

4.1 Determining the length of the LDed constituent

4.1.1 *Dependent variable: Length of the LDed constituent*

In line with earlier studies, the size (or length) of constituents is here used as a metric for measuring structural complexity (cf. Wasow 1997: 81; Yaruss 1999: 330). Arnold et al. (2000: 35), for instance, maintain that the degree of structural complexity exhibited by a constituent can be "measured as the relative length of the […] constituents, in terms of the number of words".

In the present study, the dependent variable "length" is set up as a continuous variable, so that a priori categorization of the data is not required. To explore what factors (or predictor variables) have an effect on the length of the LDed constituents,

---

8. Hilpert and Gries (2009: 390) suggest the use of Kendall's τ, which is a correlation statistic, to find out whether an increase in one ordinal or continuous variable corresponds to an increase in a second variable. Correlations can be both positive and negative, so that increases in one variable may also systematically relate to decreases in another (see Hilpert 2013: 30).

a linear regression model (Baayen 2008: 84; Gries 2013: 261; Levshina 2015: 139) is employed. Multiple linear regression modeling is here employed for "effect estimation" (Harrell 2015: 98ff). This means that, instead of trying to accomplish a minimal adequate model, the statistical model includes all theoretically relevant factors (see Baayen 2013: 348; Levshina 2015: 149). In order to build a reliable model, residuals were plotted and 37 outliers were discarded: "length" was limited to 25 words. The model is computed on a total of 750 tokens. Logarithmic values are employed for "length" in order to accommodate the fairly large range of variation in the complexity of the LDed constituents. The linear model includes "period" as an independent variable; however, in several finer counts and statistical analyses offered in Section 5.3, data from EModE and LModE is conflated: instances in the latter period are scarce (see Figure 1, Section 5.1), and the results for the EModE data alone are not significantly different from the aggregate.

4.1.2   Predictor variables

Intervening material
To the best of the author's knowledge, previous accounts of English LD have not yet examined the effect of the amount of intervening material between a LDed item and the main clause containing the resumptive item on the length of the LDed constituent. Intervening material, indicated by means of underlining in the examples, may have the shape of an aside (11), a relative clause (12), an *ing*-clause (13), or a subordinate clause (which may actually include reference to the dislocated referent) (14).

(11)   And *the woods*, <u>my man sayes</u>, he dar give a thowsand marke for **them**,
(*PPCEME*, letter, E1)

(12)   *That persone*, <u>whoe is to receave a good torne and doth disdaine to seke the same</u>, I accompt **hym** worthie to go without it.   (*PCEEC*, letter, E2)

(13)   *Simon*, <u>beinge a child of six yers old</u>, his father loved **him** above all the reste, but his mother nor brethren loved him not.   (*PPCEME*, autobiography, E2)

(14)   And *those that will be ignorant*, <u>if they find so great felicity in it</u>, let **them** be so.   (*PCEEC*, letter, E3)

Intervening material was coded (or measured) in terms of number of words, with a limit of 50 words. Raw values (instead of logarithmic) were employed in the model since tokens lacking intervening material have a value of zero.

Augmentation

The presence ("augmented") or absence ("unaugmented") of a grammatical or "procedural" item (see Traugott & Trousdale 2013) preceding the LDed constituent was coded in order to test whether augmentation can be associated with more complex LDed items (see Section 3.1). As will be seen in Section 4.2, augmentation will also be employed as a predictor variable of speech-related genres (see also Section 3.2). Some of the most typical procedural items preceding LDed strings in the data are conjunctions such as *and* (see (11) and (14) above)*, for* (15), *but* (16), *so* (see (7) above), or the complementizer *that* (17). Less frequently, the LDed string may be preceded by adverbials such as *then*, *therefore*, or *though* (among others).

(15) For *suche thinges as my good sister wryteth for* she shall receyve **them** all together with a letter the next tyme. (*PPCEME*, letter, E2)

(16) but *the moment that a master, from a sense of duty, luxuriates in corrections which do not benefit the boy*, **that moment** the master is ceasing to do his duty. (*PPCMBE*, educational treatise, L3)

(17) And that *anie Townes that will be content to geiue themselfes to the Kinge* it shall be lawefull for him to receiue **them**, so as he doe not use therein act of hostilitie. (*PCEEC*, letter, E2)

Punctuation

LDed items are separated from the main clause by a suprasegmental pause in speech (Dik 1997: 393, Ford et al. 2003: 138). In general, LDed constituents in the current data set mostly feature a following punctuation mark (usually a comma, sometimes a colon, and seldom a hyphen) reflecting the suprasegmental pause; however, the expected punctuation mark is occasionally omitted. Compare examples (18) and (19):

(18) *The messenger that brought them* I herde **hym** sey he departede from the Kynge at the Tower of London vpon Frydey last past. (*PCEEC*, letter, E1)

(19) *Tom Carrew*, **he** was an honest man which makes me remember him amongst his betters. (*PPCEME*, letter, E2)

The variable "punctuation" includes two levels: "presence" or "absence". The factor "presence or absence of a punctuation mark following the LDed constituent" was coded for in order to determine whether there is a correlation between the presence of punctuation marks and longer LDed items. As will be seen in Section 4.2, it also serves as a predictor variable for speech-related or informal genres.

Form of the resumptive

This variable ("resump.form") comprises three levels. Firstly, most previous studies and reference grammars consider personal pronouns as the prototypical resumptive in LD (Lambrecht 1994: 181; Acuña-Fariña 1996: 149; Prince 1997, 1998). These tokens have been labeled "pro" (an example is (20)). Secondly, full NPs are also often cited as resumptive items that trigger the same functional effects (Geluykens 1993; Prince 1997: 136; Van Riemsdijk 1997). These tokens (e.g. (21)) were coded as "NP".

(20)   *The people which are dead*, when they are tired of staying in the bush, then **they** come for one of their people which they like.   (*PPCMBE*, travelogue, L3)

(21)   and therefor *that Charge which the Queene hath given me for Wales in generall*, **the same** I must give you in particular for that Syrcuite.

(*PPCEME*, biography, E2)

Lastly, tokens were classified as "demonstrative" when the resumptive slot holds a demonstrative pronoun, as in (22).

(22)   On the whole, however, you are not to take gloomy views, for there is nothing to mourn at, to despair at: *a serious cheerfulness*; **that** is the right mood in this as in all cases.   (*PPCEME*, history, L2)

The level "demonstrative" captures a functional nuance that was mainly active before 1500, namely whether or not a resumptive item can be said to have contrastive power and, thus, be capable of licensing CLD. As Los & Komen (2012: 896) have shown, demonstrative resumptives constitute a reliable trace of the steep decline of CLD since OE. Distinguishing this third level of the factor "resump.form" should help us determine (i) the extent to which CLD may still be present in the ModE data and (ii) whether such demonstrative resumptives correlate with more or less complex LDed constituents.

Position of the resumptive

The variable "resump.position" includes two levels: "subject-first" and "other". Examples (23) and (24) illustrate a pronominal and demonstrative in "subject-first" position, while (25) is an instance classified as "other".

(23)   *The other which ys Jhonne Lynton* **he** hathe £12 a yeare lande yf he canne keape yt, wheare of there ys some dought.   (*PCEEC*, letter, E2)

(24)   *The fleeting thinges as ayre & water*, **these** easely be departed,

(*PPCEME*, philosophy, E2)

(25)   And *he that taketh one oute of saintuary to dooe hym good*, I saye plainely that **he** breaketh no saintuary.   (*PPCEME*, history, E1)

Resumptive items simplify processing, but not all of them might be expected to do so equally: pronominal resumptive items in subject-first (or preverbal subject) position are expected to corefer with longer items since they will process the detached item more rapidly and at a lower memory cost (see Section 2.2).

Genre cluster

Following the typology in Culpeper and Kytö (2010: 16–18), the genres included in the corpora were classified into four different categories: (i) writing-related genres, represented by biography, educational treatise, handbook, history, law, philosophy, science, and travelogue; (ii) mixed genres such as prose fiction and trial proceedings; (iii) speech-purposed registers, like drama and sermons; and (iv) speech-like genres, such as diaries and letters. Table 1 specifies the classification applied.

Table 1. Genres organized into clusters according to medium and setting

	Level of formality	**Informal**	**Formal**
Medium			
Speech-related	speech-like	Letters-private Diaries	Letters-official
	speech-purposed	Drama, comedy	Sermons
Mixed			Prose fiction
			Trial proceedings
Writing-related			Biography
			Educational treatise
			Handbook
			History
			Law
			Philosophy
			Science
			Travelogue

LDed constituents have been reported to be shorter in speech than in writing (see Section 1). It can therefore be hypothesized that the writing-related genres will predict longer LDed items and that speech-related genres will be a factor facilitating shorter items. In the linear model (see Table 2 below), the reference level for this variable is "mixed".

Period

For this variable, tokens were coded on the basis of the labels specified in the corpora documentation, namely "E1" (1500–1569), "E2" (1570–1639), "E3" (1640–1710), "L1" (1700–1769), "L2" (1770–1839), "L3" (1840–1914). Complexity in LD has seldom been examined in terms of the factor time or period. Pérez-Guerra (1999: 226) gives mean lengths of 8.9 words in late Middle English, 8.8 in Early Modern English, and 7.4 in Present Day English. Tizón-Couto (2015: 45) reports mean lengths of 9 words in EModE, 7.1 words in LModE, and 6.3 words in PDE. In line with the mean values reported in previous studies, length of the LDed constituent is also expected to decrease across time.

4.2   Exploring the link between speech-related genres and complexity of the LDed constituent

In order to explore the correlation between less complex (short and/or unaugmented) LDed constituents and speech-related genres, "genre cluster" was selected as the dependent variable. Predictor variables are "period", "augmentation", "punctuation", and "type of LDed string". This last variable classifies tokens into several levels: "LD" when the LDed string conforms to the prototypical LD pattern exemplified by (2) and (3) in Section 1; "listLD" when the LDed string contains a list of LDed NPs, as in (4) above; the three remaining types ("LDwh", "LDing", and "whLD") involve a relative clause postmodifying the LDed NP (5), an LDed Absolute Construction (6), or an LDed Free Relative clause (7).

   Given that "genre cluster" is a categorical variable with four levels (see Table 1), the effects of predictor variables are assessed by means of univariate statistical tests: Kendall's Correlation tests for "period" and Pearson's chi-squared tests for the other predictors, namely "augmentation", "punctuation" (see Section 4.1 for descriptions), and "type of LDed string".

5.   Results

Section 5.1 offers a general overview of the distribution of LDed strings from ME onwards in the *Penn-Helsinki corpora*, thus providing a backdrop for the ensuing discussion of the more specific results for ModE. Section 5.2 presents the results of the study of the factors predicting the complexity of LDed constituents; complexity is investigated by means of a multiple linear regression model that takes the "length" of the LDed constituent (i.e. a measure of complexity) as the dependent variable. Section 5.3 investigates the alleged association between speech-related genres and

(short and unaugmented) LDed constituents; univariate statistical tests are employed here to test the effects of predictors on "genre cluster".

5.1 The diachrony of LD: An overview

The overall data confirm the previously reported loss of LDed strings from ME onwards (Kendall's Correlation test = −1, $p < 0.001$). Figure 1, which portrays the change in frequency by genre, shows that biblical texts are the only genre where LDed strings do not decline. This can be attributed to the very special syntax of these biblical texts, even before translation into English.[9] This result also justifies the exclusion of tokens from the Bible from the investigated dataset, thus allowing a finer analysis of ModE usage data in Sections 5.2 and 5.3.
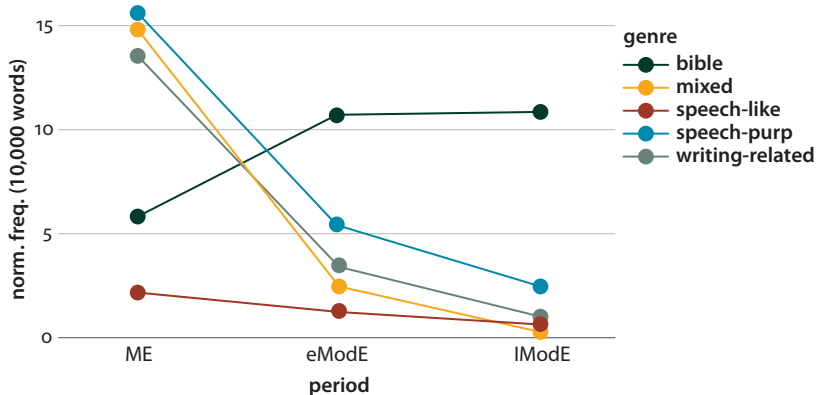
Figure 1. Historical evolution of LDed strings by genre cluster (normalized per 10,000 words)

---

9. In the Penn-Helsinki data, the Bible is the written genre with the highest density of LDed strings (0.79 per 1,000 words). Such a high figure, in comparison with genres where paratactic syntax can be expected (drama: 0.35 or sermon: 0.27), is explained by the fact that both the Greek and Hebrew original biblical texts display many examples of left detachments and adverbial frame-setting constructions (see van der Merwe & Talstra 2003; Runge 2010: 287–312; Holmstedt & Jones 2014). Early and Late Modern versions of the (English) Bible constitute fairly literal translations of the Greek and Hebrew original; accordingly, the archaic word order was probably preserved.

The overall decrease of tokens continues from EModE (2.45 per 10,000 words) into LModE (1.32 per 10,000 words).[10] In this period, when SVO word order becomes fixed, LDed strings display clear genre variation. The analysis of this genre variation is presented in Section 5.3.

As regards the evolution of CLD alone, the current dataset includes 144 tokens (out of the 787 extracted after excluding the Bible) that feature a demonstrative resumptive. Thus, only 16.4% of the tokens might be confidently claimed to express the contrastive function in the ModE data. Fully in line with the general decrease reported for CLD since OE in Los (2009) and Los & Komen (2012), the results suggest that the contrastive variant becomes an extremely low frequency construction in ModE, even in comparison with the already scarce LD. Furthermore, demonstrative resumptives occupy the preverbal subject (or subject-first) position in 107 tokens of the 144 that positively qualify as CLD (74.3%). This particular result backs up Los & Komen's (2012: 896) claim that "CLD remained [only] available for dislocates with a resumptive subject".

5.2   Results from the linear model: Complexity of the LDed constituent

This section reports on the factors (significantly) predicting the length of LDed constituents, as they emerge from the linear regression model summarized in Table 2. The model was built with seven fixed predictor terms (see Section 4.1.2) by means of the "lm" ("fitting linear models") function in R, and then checked for multicollinearity by means of the "vif" ("variance inflation factors") function included in the "car" package ('companion to applied regression'; Fox & Weisberg 2011). The generalized variance inflation factors (GVIFs) generated were are all well below the safety threshold of 4 suggested by Fox and Weisberg (2011: 327) for linear regression modeling. Note that the R-squared values for the model are low due to the high amount of variation between tokens in terms of the "length" of the LDed item (i.e. the dependent variable).

---

10. This decline might be linked to the establishment of the syntactic and orthographical bases of the sentence. Robinson (1996), for instance, suggests that "a conceptual change from the 'period', defined aurally and rhetorically, to the 'sentence', defined visually and syntactically, should be dated to the mid-to-late seventeenth century" (in Lennard 1995: 67).

Table 2. Multiple linear regression model, with "length" of the LDed constituent as the dependent variable. Note that a positive estimate implies longer LDed items

Residuals:	Min	1Q	Median	3Q	Max
	−2.229	−0.459	0.028	0.513	2.042

Coefficients:					
	Estimate	Sth. Error	t value	Pr(> |t|)	signif.
(Intercept)	1.615	0.132	12.196	< 0.001	***
intervening.material					
(continuous)	−0.026	0.003	−7.706	< 0.001	***
augmentation (ref. level = 'unaugmented')					
'augmented'	0.218	0.053	4.043	< 0.001	***
punctuation (ref.level = 'unpunctuated')					
'punctuated'	0.272	0.061	4.473	< 0.001	***
resump.form (ref.level = 'demonstrative')					
'NP'	0.046	0.121	0.38	0.704	
'pro'	−0.338	0.079	−4.277	< 0.001	***
resump.position (ref.level = 'other')					
'subject1st'	−0.102	0.053	−1.898	0.058	.
genre.cluster (ref.level = 'mixed')					
'speech-like'	0.240	0.097	2.45	0.014	*
'speech-purposed'	−0.025	0.108	−0.23	0.818	
'writing-related'	0.203	0.097	2.087	0.037	*
period (ref.level = 'E1')					
'E2'	−0.126	0.065	−1.938	0.053	.
'E3'	0.051	0.076	0.66	0.509	
'L1'	−0.411	0.196	−2.093	0.036	*
'L2'	−0.414	0.141	−2.926	0.004	**
'L3'	0.411	0.168	2.43	0.015	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.7097 on 735 degrees of freedom
Multiple R-squared: 0.2015, Adjusted R-squared: 0.1863
F-statistic: 13.25 on 14 and 735 DF, p-value: < 0.001

Six of the factors considered in the linear model significantly determine the variation: (i) "intervening.material", (ii) "augmentation, (iii) "punctuation", (iv) "resump.

form", (v) "genre.cluster", and (vi) "period". The variable "resump.position" only approaches borderline statistical significance ($p < 0.06$).

The factor "intervening.material" predicts that the longer the intervening material, the shorter is the LDed item. However, this effect must be interpreted with caution: what the model truly suggests is that LDed items frequently lack directly following material. When tokens that feature no intervening material are excluded, the correlation between the numeric predictor and dependent variables does not reach significance. This is illustrated by means of a scatter plot displaying the logarithmic values for both "length" and "intervening.material" (Figure 2), which is based on the 230 tokens that feature at least one intervening constituent.
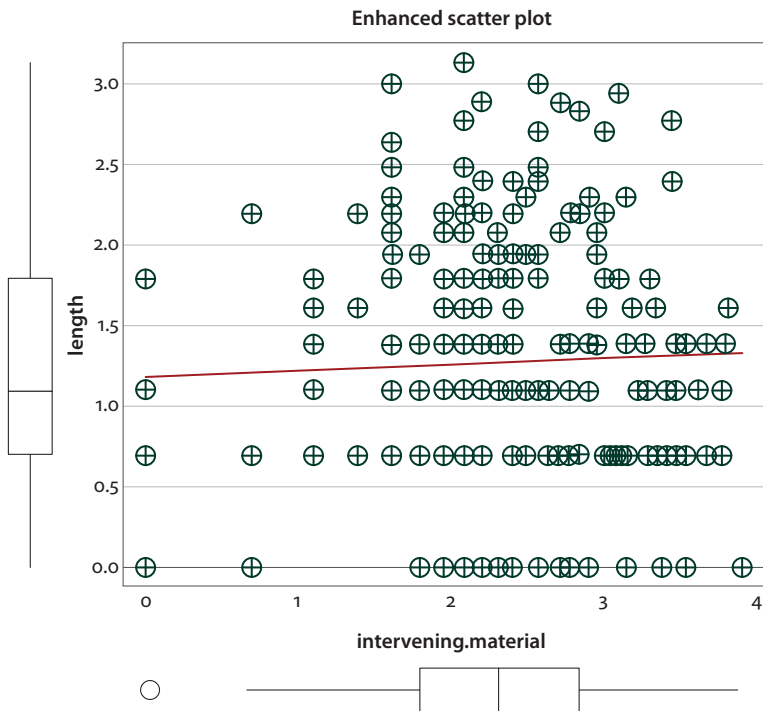
Figure 2. Scatter plot on the correlation between the logarithmic values for both "length" and "intervening.material" (including only tokens that feature intervening material)

The linear model also indicates that LDed constituents tend to be longer when they are preceded by a conjunction, preposition, or adverb ("augmentation_augmented") and when they feature a comma signaling the suprasegmental pause between the detached item and the main clause ("punctuation_punctuated"). The effect plots in Figure 3 illustrate these two significant correlations, respectively.
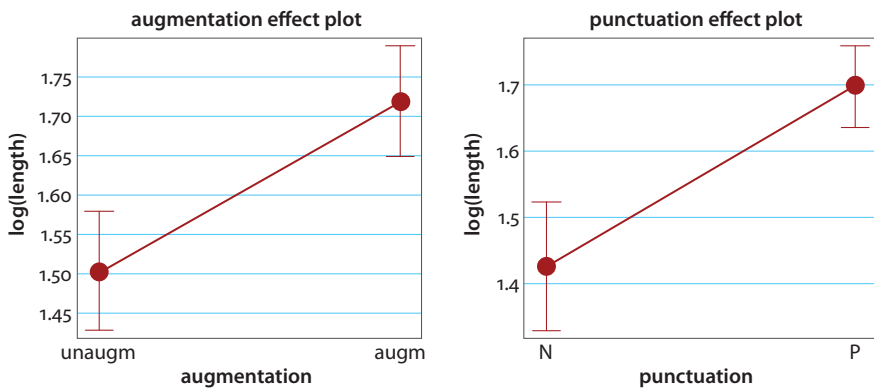
Figure 3. Effect plots for "augmentation" (left) and "punctuation" (right); y-axis = log ("length")

The statistical model signals a tendency for LDed constituents to be shorter when resumed by means of a pronominal resumptive item ("resump.form_pro"). Further, although the p-value for "resump.position_subject1st" does not reach the .05 threshold, it suggests a tendency for subject-first resumptives to corefer with shorter LDed items. Both results, illustrated by means of the two effect plots in Figure 4, defy the most logical link (a priori) in terms of processing, namely that longer items should corefer with a nearby pronominal element. Only demonstrative resumptives clearly correlate with longer items (see the first effect plot in Figure 4), which could have to do with their specific contrastive function rather than with general issues of processing.
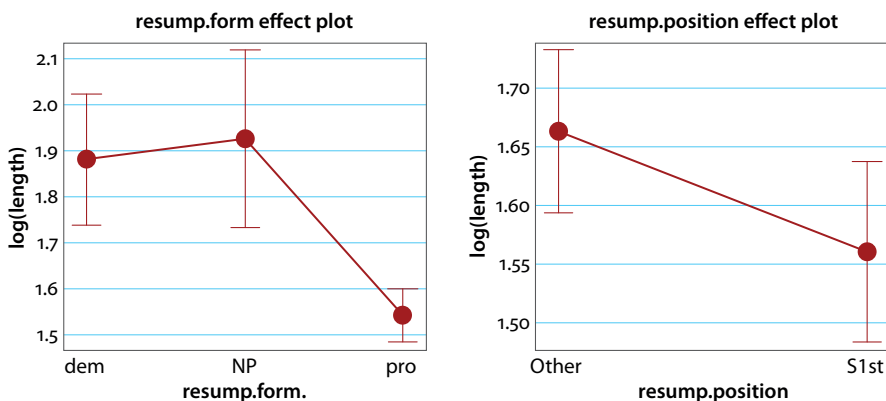
Figure 4. Effect plots for "form of resumptive" (left) and "position of resumptive" (right); y-axis = log ("length")

The model furthermore predicts a significant effect of genre. The longest LDed items significantly correlate with "writing-related" genres, as expected. There is, however, also a significant correlation between "speech-like" genres and longer items; this is contrary to expectations, as shorter LDed constituents would be anticipated in texts that more closely resemble speech. A very important point is that, in contrast with "speech-like" and "writing-related" texts, both "mixed" (the model's reference level for this variable) and "speech-purposed" genres frequently feature shorter items: note the lack of an asterisk for "speech-purposed" in the model, as "mixed" is the reference level. Furthermore, a separate calculation of the mean lengths for the LDed strings in these two clusters (5.4 words for "mixed" and 5.5 for "speech-purposed") indicates that both rank closer to the average length of spoken LD (roughly 5 to 5.5 words; see Section 1). The relevant distinctions are summarized in the first plot in Figure 5.
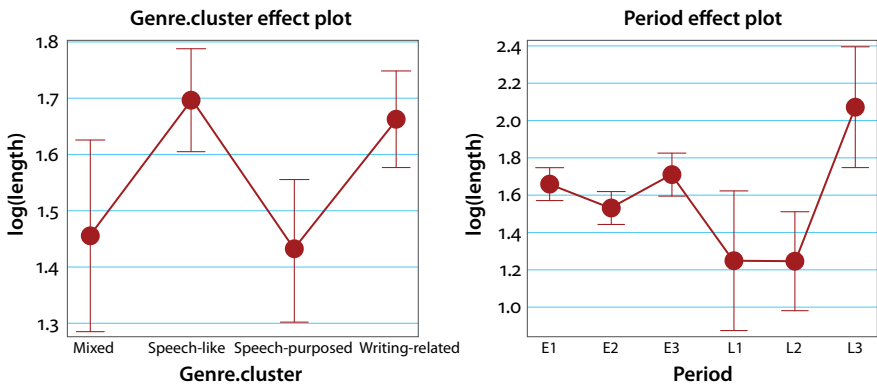
Figure 5. Effect plots for "genre.cluster" (left) and "period" (right); y-axis = log ("length")

Finally, "period" also has an effect on "length": there is a tendency for the complexity of the LDed constituent to decrease in time, except for the last subperiod in LModE. The second effect plot in Figure 5 illustrates that LDed constituents are shorter for L1 and L2; however, L3 significantly breaks up the LModE trend. Because LDed strings are low-frequent in the LModE data (see Figure 1), it is not possible to ascertain whether the results for L1 and L2 might be somehow deflated due to chance or, conversely, whether it is the average for L3 that is inflated. The latter is more likely because L3 contains the lowest number of examples. Clearly, then, more data from the LModE period are required in order to confirm the general diachronic tendency towards shorter LDed items. In this respect, the validation of such declining trend appears to constitute a challenging endeavor given that LD visibly becomes a low frequency construction during LModE (see also Tizón-Couto 2015).

5.3   Results from univariate statistical tests: Orality and genre

This section explores the potential link between the LDed NPs and the speech-related texts in the ModE corpora investigated. It begins by statistically evaluating the diachronic development of LDed strings (comprising LDed NPs) within each separate genre cluster across the ModE subperiods investigated. Kendall's correlation test is employed as a means to assess the extent to which these LDed items, which have been strongly associated with contemporary speech, are diachronically more or less resilient in genre clusters that have been claimed to be more or less "oral" (cf. Culpeper & Kÿto 2010). This section also reports the results of a series of chi-squared tests that take "genre cluster" as their dependent variable. A number of predictors that embody the features typically associated with LDed NPs in speech (other than "length", for which see 5.2 above) are employed to assess variation: "augmentation", "punctuation", and "type of LD". The results are presented in the shape of mosaic plots with residual-based shadings (Meyer et al. 2006; Zeileis et al. 2007; Meyer et al. 2016).

Figure 6 shows the normalized frequencies of LDed items tagged as NP (illustrated in (1)–(5)) in the *Penn Parsed corpora* for each genre cluster from EModE to LModE.

Kendall's tau test indicates that the decrease across the six subperiods in Figure 6 is only significant for the "writing-related" and the "speech-like" clusters (Kendall's tau = −0.73, $p = 0.05$ and −1, $p < 0.01$, respectively). Thus, frequency of LDed strings declines significantly in "speech-like" texts, like letters and diaries, where a more profound influence of orality (on the part of the writer) might be expected (cf. Culpeper & Kytö 2010). This is a finding that runs counter to the expectation that LDed NPs might constitute markers of unintended orality or informality. Note, also, that the overall relative frequency of tokens, including both EModE and LModE, in speech-like texts (1.23 per 10,000 words) is lower than in "speech-purposed" (4.28), "mixed" (1.89), or "writing-related" (2.31) genres. In contrast, LDed string frequency does not significantly decrease for genres where writers may deliberately attempt to reproduce orality, namely both "speech-purposed" genres like drama and sermon and "mixed" genres such as fiction and trial proceedings.

Significant correlations hold between "genre cluster" and each of the three predictors considered here in order to explore the extent to which the features of LDed items might help establish the degree of orality of the genre clusters: (i) LDed NPs are more frequently "unaugmented" in speech-purposed texts and more frequently "augmented" in writing-related texts (see Figure 7); (ii) punctuation marks are often omitted after the LDed constituent in speech-like texts (see Figure 8); and (iii) LDed NPs featured in speech-purposed texts are more
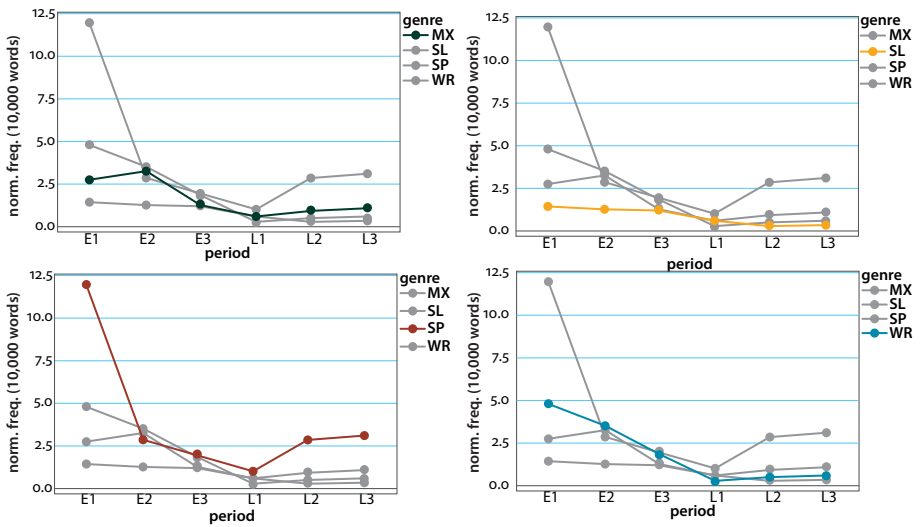
Figure 6. Historical evolution of LDed strings by genre
*Genre clusters: Mixed (n = 71), Speech-Like (n = 269), Speech-Purposed (n = 141),
Writing-Related (n = 306)*
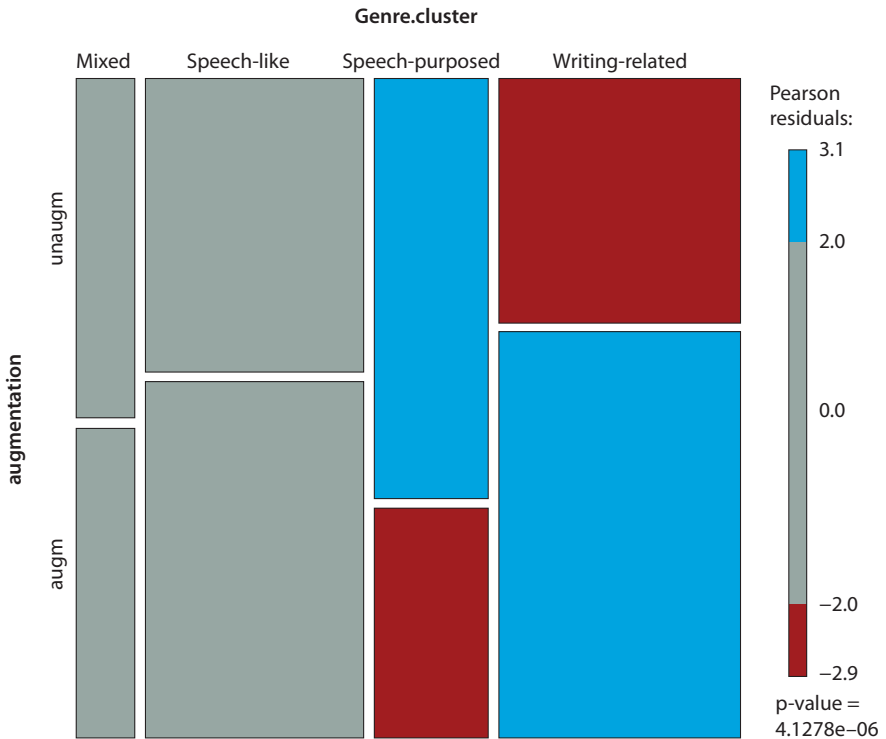
Figure 7. Mosaic plot for "genre cluster" and "augmentation"
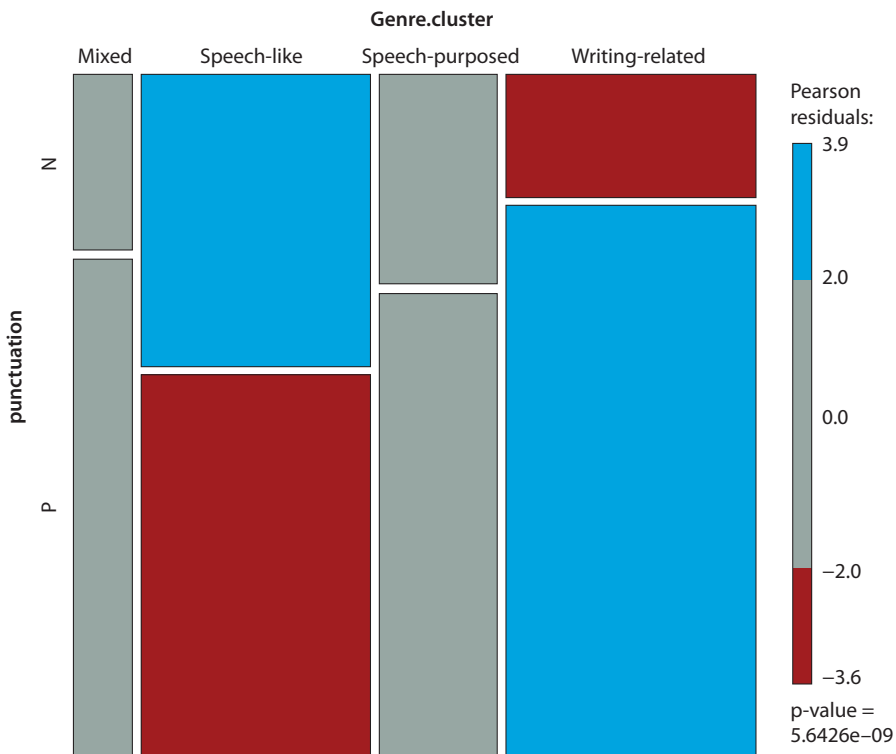
Figure 8. Mosaic plot for "genre cluster" and "punctuation"

frequently similar to the construction that is typically cited as LD in reference grammars (see Figure 10).[11]

In Figures 7 and 8, the colors attribute a significant value of the Pearson residual (i.e. the individual contribution to the Pearson chi-squared statistic) to the cells (or combinations of levels). The legend for the interpretation of the residuals is presented at the right of the plot. More specifically, blue means there are more observations in that cell than would be expected under the null model (independence). Red means there are fewer observations than would have been expected. In Figure 7, the distributional difference between tokens in "speech-purposed" and "writing-related" genres is reflected in their contrasting shading patterns: drama and sermons more frequently include "unaugmented" LDed NPs. This means that the unaugmented LDed items that have been associated with (contemporary)

---

11. The plots in Figures 7 to 10 are computed on the same reduced dataset (750 tokens) as the linear model in Table 2. The raw figures for the "Genre.cluster" variable in these figures are the following: 69 tokens are "mixed", 261 are "speech-like", 135 tokens are "speech-purposed", and 285 tokens are "writing-related".
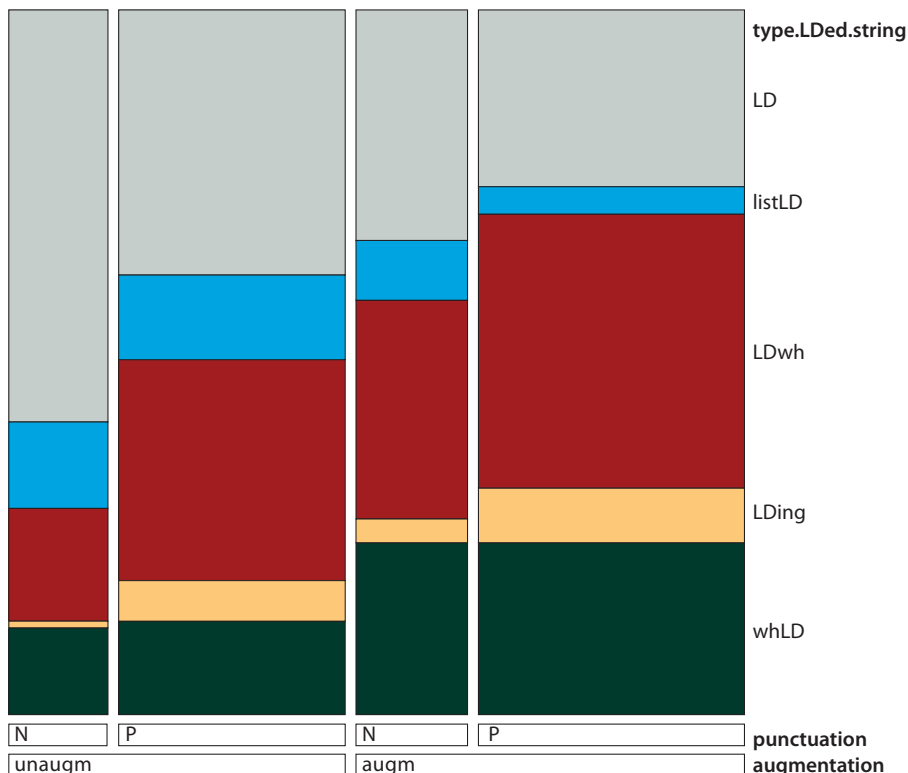
Figure 9. Double-decker plot for "type of LDed string", "punctuation" and "augmentation" "type of LDed string"

speech in previous studies (cf. Ono & Thompson 1994; Hidalgo 2000; Fox et al. 2003) are also likely to have been considered fit to reproduce or mock speech by ModE authors. Figure 8 indicates that the signaling of the suprasegmental pause that naturally follows a LDed item in speech (cf. Dik 1997: 393; Ford et al. 2003: 138) is less consistently marked in "speech-like" genres, in stark contrast with the "writing-related" cluster. This result might be attributed to tighter editing rules applied in the more formal ("writing-related") texts or to the Latin-like periodic style that was imitated in certain "writing-related" genres until the 19th century (cf. Görlach 1999: 88; 2001: 128). Another possible interpretation is that writers actually used punctuation as a result of their unconscious awareness that more complex LDed items might imply processing/reading issues. This is a rather speculative (and unlikely) analysis, however, given that "speech-like" texts actually include significantly longer LDed items in comparison with the "speech-purposed" and "mixed" clusters. Thus, the statistical distinctions in Figure 9 possibly have more to do with differing stylistic preferences and editorial practices across genres (cf. van Kemenade & Los 2013: 221).

Figure 9 is meant to facilitate the interpretation of the correlation tested in Figure 10 between "genre cluster" and "type of LDed string". Figure 9 is a double-decker plot (Meyer et al. 2006) which shows the distribution of "augmentation" and "punctuation" across the types of LDed string proposed in this study. Figure 9 conveys that the types more strongly associated with speech in earlier literature, namely "LD" and "listLD", tend to correlate in the ModE data with lacking augmentation preceding the LDed item and lacking punctuation following that item. In contrast, items that could be analyzed as clausal (i.e. "LDing" and "whLD") are most usually preceded by a grammatical word ("augmented") and separated from the ensuing clause by a punctuation mark. These results suggest that the "LD" and "listLD" types are the most reliable as potential signposts of orality or informality in historical texts.
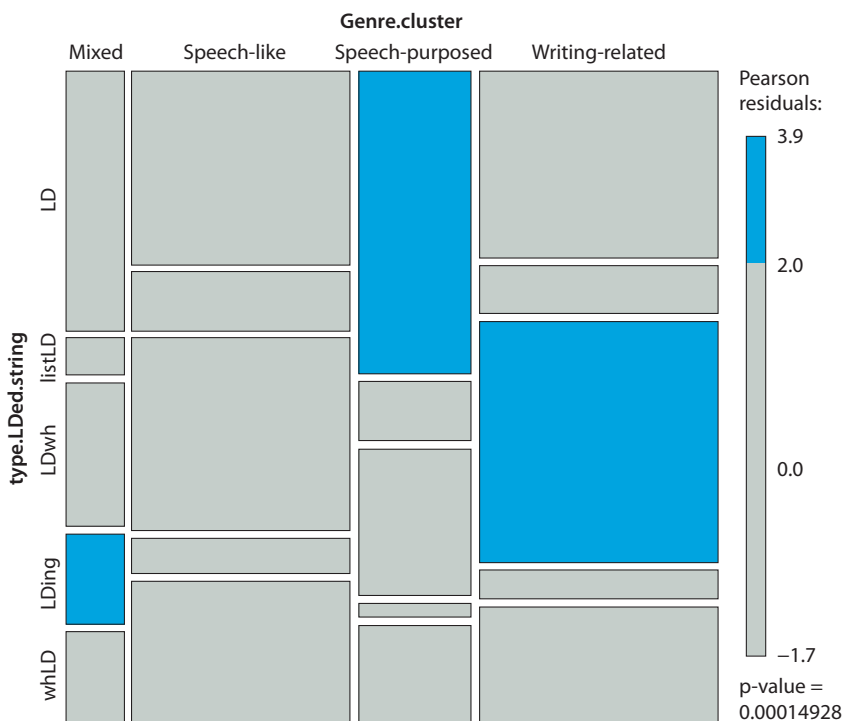
Figure 10. Mosaic plot for "genre cluster" and "type of LDed string"

The blue areas in the mosaic plot in Figure 10 suggest that the "speech-purposed" cluster is the only one that includes a significantly higher rate of oral-like LDed items (i.e. "LD"), while the "writing-related" cluster is characterized by LDed items which feature heavy post-modification in the shape of a relative clause (i.e. "LDwh"). Lastly, the "mixed" cluster includes a more significant number of items which could be analyzed as absolutes.

In sum, the LDed items that the literature has outlined as characteristic of speech (unaugmented, shorter, and less heavily postmodified) are proportionally more frequent in "speech-purposed" texts than in "speech-like" texts. This means that LDed strings were likely employed in order to intentionally reproduce conversation, or speech in general, but are unlikely to constitute traces of truly unintended orality on the side of the writer. It might also suggest that letters and diaries could be less representative of orality/speech than previously argued, at least when paratactic strategies such as LD are in focus. However, this much wider speculation clearly requires further research.

6.   Discussion and concluding remarks

A general finding of this paper is that the overall significant decrease of LDed strings in ModE is in line with the stark decrease previously reported for LD from OE onwards. In addition, the number of tokens that can be classified as carrying out a contrastive function (CLD) is very low and tightly restricted to the preverbal (or subject-first) position.

The first specific goal of the study was to explore the factors that determine the complexity of LDed constituents in ModE. In the current dataset, variation in the length of the LDed item is partly determined by principles of processing complexity: firstly, there seems to be a certain balance concerning the degree of complexity allowed before a main clause in that a long LDed constituent prevents the construction from featuring intervening material. However, the results do not allow for the generalization that the amount of intervening material conditions length, as there is no correlation between the word counts of the intervening material and the length of the LDed constituent. Secondly, a preceding grammatical word aids the parsing of the referential link to an LDed constituent that is more heavily postmodified: greater syntactic weight is associated with greater processing complexity, which leads to a more explicit (i.e. "augmented") LDed constituent.

The fact that longer LDed items are usually followed by a punctuation mark or an intervening segment may be linked to the periodic nature of particular writing-related texts, which imitated Latin models. Such periodic style has been claimed to endure well into the nineteenth century (Görlach 1999: 88; 2001: 128). Besides possibly indicating stylistic preferences that go hand in hand with writing-related genres and a higher degree of formality, the correlation of punctuation marks with lengthier LDed items may also indicate an acknowledgement, on the side of the writer, of a higher degree of processing complexity. Nonetheless, more evidence would clearly be needed in order to confirm this conjecture, as well as a better understanding of punctuation conventions in the period.

No confirmation has been found in our dataset for the hypothesized proximity constraint within the construction that more complex LDed items should predominantly corefer with subject-first pronominal resumptives in the ensuing sentence. Therefore, a nearby pronominal resumptive item does not appear to constitute an accommodation strategy to deal with longer LDed constituents, at least in writing.

The second goal of the study was to explore the connection between LDed items and orality and between less complex LDed items and orality. From what can be gleaned from the data at hand, LDed strings in speech-like genres follow similar patterns to those found in writing-related texts in terms of their diachronic development and distinctive features. Their significant decrease in LModE, together with their formal features ("augmented" and more heavily postmodified), suggests that LDed items did not generally mark genuine orality in ModE and, thus, do not appear to constitute unplanned devices that illustrate the higher degree of communicative immediacy that Culpeper and Kytö (2010), for instance, claim for letters and diaries. The clearest sign of informality seems to be that punctuation marks can be easily omitted for the speech-like set. Contrary to what can be reported for speech-like texts, tokens in speech-purposed and mixed texts show characteristics that bring them closer to the LD construction that has been often cited as typical of contemporary speech: they are unaugmented and do not include heavy postmodification. In addition, they do not undergo a significant diachronic decline after the fixation of SVO. Therefore, the prediction that LDed NPs are employed as a trait of orality is only borne out for particular genres that deliberately reproduce conversation or speech (cf. Geluykens 1992, 1993; Leonarduzzi & Herry 2005).

Overall, the results of this study provide new insights into (i) the declining diachronic trend of LDed strings that went hand in hand with the fixation of SVO syntax, (ii) the accommodation strategies that might facilitate the resumption of more complex LDed items (e.g. an extant previous procedural item or the presence of a following punctuation mark), and (iii) the ability of LDed strings to intentionally reproduce, but not inherently represent, orality in ModE speech-related texts. As for further research, the most logical step forward is to achieve more detail as regards matters of both complexity and orality in speech-like data alone (cf. Tizón-Couto 2016).

Acknowledgements

References

Acuña-Fariña, Juan Carlos. 1996. *The Puzzle of Apposition: On So-called Appositive Structures in English*. Santiago de Compostela: Universidade de Santiago de Compostela (Servicio de Publicacións e Intercambio Científico).

Anagnostopoulou, Elena. 1997. Clitic left dislocation and contrastive left dislocation. In Anagnostopoulou, van Riemsdijk & Zwarts (eds), 151–192.

Anagnostopoulou, Elena, van Riemsdijk, Henk & Zwarts, Frans (eds). 1997. *Materials on Left Dislocation* [Linguistik Aktuell/Linguistics Today 14]. Amsterdam: John Benjamins. doi:10.1075/la.14

Andrews, Avery D. 2007. Relative clauses. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, 2nd edn., Timothy Shopen (ed.), 206–236. Cambridge: CUP. doi:10.1017/CBO9780511619434.004

Arnold, Jennifer E., Wasow, Thomas, Losongco, Anthony & Ginstrom, Ryan. 2000. Heaviness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76(1): 28–55. doi:10.1353/lan.2000.0045

Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: CUP. doi:10.1017/CBO9780511801686

Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Birner, Betty & Ward, Gregory. 2002. Information packaging. In *The Cambridge Grammar of the English Language*, Rodney Huddleston & Geoffrey Pullum (eds), 1363–1447. Cambridge: CUP.

Culpeper, Jonathan & Kytö, Merja. 2010. *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: CUP.

Diessel, Holger. 2005. Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43: 449–470. doi:10.1515/ling.2005.43.3.449

Dik, Simon. 1997. *The Theory of Functional Grammar, Part II: Complex and Derived Constructions*, edited by Kees Hengeveld. Berlin: Mouton de Gruyter. doi:10.1515/9783110218374

Dons, Ute. 2004. *Descriptive Adequacy of Early Modern English Grammars*. Berlin: Mouton de Gruyter. doi:10.1515/9783110906042

Fischer, Olga. 1992. Syntax. In *The Cambridge History of the English Language, Vol. 2: 1066–1476*, Norman Blake (ed.), 207–408. Cambridge: CUP. doi:10.1017/CHOL9780521264754.005

Fonteyn, Lauren & van de Pol, Nikki. 2016. Divide and conquer: The formation and functional dynamics of the Modern English *ing*-clause network. *English Language and Linguistics* 20(2): 185–219.

Ford, Cecilia, Fox, Barbara & Thompson, Sandra A. 2003. Social interaction and grammar. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Vol. 2., Michael Tomasello (ed.), 119–143. Mahwah NJ: Lawrence Erlbaum Associates.

Fox, John & Weisberg, Sanford. 2011. *An R Companion to Applied Regression*, 2nd edn. Thousand Oaks CA: Sage.

Geluykens, Ronald. 1992. *From Discourse Process to Grammatical Construction: On Left Dislocation in English* [Studies in Discourse and Grammar 1]. Amsterdam: John Benjamins. doi:10.1075/sidag.1

Geluykens, Ronald. 1993. Syntactic, semantic and interactional prototypes: The case of left-dislocation. In *Conceptualizations and Mental Processing in Language*, Richard A. Geiger & Brygida Rudzka-Ostyn (eds), 709–730. Berlin: Mouton de Gruyter. doi:10.1515/9783110857108.709

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1): 1–76. doi:10.1016/S0010-0277(98)00034-1

Givón, Talmy. 1979. *On Understanding Grammar*. New York NY: Academic Press.

Görlach, Manfred. 1999. *English in Nineteenth-Century England: An Introduction*. Cambridge: CUP. doi:10.1017/CBO9780511627828

Görlach, Manfred. 2001. *Eighteenth-Century English*. Heidelberg: Universitätsverlag C. Winter.

Gregory, Michelle & Michaelis, Laura. 2001. Topicalization and left-dislocation: A functional opposition revisited. *Journal of Pragmatics* 33: 1665–1706. doi:10.1016/S0378-2166(00)00063-1

Gries, Stefan T. 2013. *Statistics for Linguistics with R*. Berlin: De Gruyter Mouton. doi:10.1515/9783110307474

Harrell, Frank E. 2015. *Regression Modeling Strategies*, 2nd edn. New York NY: Springer. doi:10.1007/978-3-319-19425-7

Hawkins, John A. 1992. Syntactic weight and information status in word order variation. *Linguistische Berichte* 4: 196–219.

Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: OUP. doi:10.1093/acprof:oso/9780199252695.001.0001

Hidalgo, Raquel. 2000. Establishing topic in conversation: A contrastive study of left dislocation in English and Spanish. In *Talk and Text: Studies on Spoken and Written Discourse*, Angela Downing Rothwell, Jesús Moya Guijarro & José I. Albentosa Hernández (eds), 137–158. Cuenca: Ediciones de la Universidad de Castilla-La Mancha.

Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: CUP. doi:10.1017/CBO9781139004206

Hilpert, Martin & Gries, Stefan T. 2009. Assessing frequency changes in multi-stage diachronic corpora: Applications for historical corpus linguistics and the study of language acquisition. *Literary and Linguistic Computing* 24(4): 385–401. doi:10.1093/llc/fqn012

Holmstedt, Robert & Jones, Andrew R. 2014. The pronoun in tripartite verbless clauses in Biblical Hebrew: Resumption for left-dislocation or pronominal copula? *Journal of Semitic Studies* 59(1): 53–89. doi:10.1093/jss/fgt035

Kortmann, Bernd. 1991. *Free Adjuncts and Absolutes in English: Problems of Control and Interpretation*. London: Routledge.

Kortmann, Bernd. 1995. Adverbial participial clauses in English. In *Converbs in Cross-linguistic Perspective*, Martin Haspelmath & Ekkehard König (eds), 189–237. Berlin: Mouton de Gruyter.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: CUP. doi:10.1017/CBO9780511620607

Lennard, John. 1995. Punctuation: And – 'Pragmatics'. In *Historical Pragmatics* [Pragmatics & Beyond New Series 35], Andreas Jucker (ed.), 65–98. Amsterdam: John Benjamins. doi:10.1075/pbns.35.08len

Leonarduzzi, Laetitia & Herry, Nadine. 2005. Les dislocations: Texts et contexts. In *Actes, Congrès de la Société des Anglicistes de l'Enseignement Supérieur, Toulouse: France*. <http://hal.archives-ouvertes.fr/hal-00484204> (30 March 2008).

Levshina, Natalia. 2015. *How to do Linguistics with R*. Amsterdam: John Benjamins. doi:10.1075/z.195

Los, Bettelou. 2009. The consequences of the loss of verb-second in English: Information structure and syntax in interaction. *English Language and Linguistics* 13(1): 97–125. doi:10.1017/S1360674308002876

Los, Bettelou & Komen, Erwin. 2012. Clefts as resolution strategies after the loss of a multifunctional first position. In *The Oxford Handbook of the History of English*, Terttu Nevalainen & Elizabeth Traugott (eds), 884–898. Oxford: OUP.

Manetta, Emily. 2007. Unexpected left dislocation: An English corpus study. *Journal of Pragmatics* 39: 1029–1035. doi:10.1016/j.pragma.2007.01.003

Meyer, David, Zeileis, Achim & Hornik, Kurt. 2006. The strucplot framework: Visualizing multi-Way contingency tables with vcd. *Journal of Statistical Software* 17(3): 1–48. doi:10.18637/jss.v017.i03

Meyer, David, Zeileis, Achim & Hornik, Kurt. 2016. vcd: Visualizing categorical data. R package version 1.4–3.

Netz, Hadar & Kuzar, Ron. 2007. Three marked theme constructions in spoken English. *Journal of Pragmatics* 39: 305–335. doi:10.1016/j.pragma.2006.04.007

Ono, Tsuyoshi & Thompson, Sandra A. 1994. Unattached NPs in English conversation. *Berkeley Linguistics Society* 20: 402–419.

PPCEME. The Penn-Helsinki Parsed Corpus of Early Modern English, 1500–1710. Department of Linguistics, University of Pennsylvania. CD-ROM, 1st edn. <www.ling.upenn.edu/hist-corpora/>

PPCMBE. The Penn Parsed Corpus of Modern British English, 1700–1914. Department of Linguistics, University of Pennsylvania. CD-ROM, 1st edn. <www.ling.upenn.edu/hist-corpora/>

PPCME2. The Penn-Helsinki Parsed Corpus of Middle English, 1150–1500. Department of Linguistics, University of Pennsylvania. CD-ROM, 2nd edn. <www.ling.upenn.edu/hist-corpora>

PCEEC. The Parsed Corpus of Early English Correspondence, 1410–1695. Compiled by the CEEC Project Team. University of York and University of Helsinki. Distributed through the Oxford Text Archive.

Pérez-Guerra, Javier. 1999. *Historical English Syntax: A Statistical Corpus Based Study on the Organisation of Early Modern English Sentences*. Munich: Lincom.

Pérez-Guerra, Javier. 2012. Discourse status and syntax in the history of English: Some explorations in topicalization, left-dislocation, and *there*-constructions. In *Information Structure and Syntactic Change in the History of English*, Anneli Meurman-Solin, María José López-Couso & Bettelou Los (eds), 121–138. New York NY: OUP. doi:10.1093/acprof:oso/9780199860210.003.0006

Pérez-Guerra, Javier & Tizón-Couto, David. 2009. On left dislocation in the recent history of English: Theory and data hand in hand. In *Dislocated Elements in Discourse: Syntactic, Semantic, and Pragmatic Perspectives*, Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (eds), 31–48. London: Routledge.

Prince, Ellen. 1997. On the functions of left-dislocation in English discourse. In *Directions in Functional Linguistics* [Studies in Language Companion 36], Akio Kamio (ed.), 117–144. Amsterdam: John Benjamins. doi:10.1075/slcs.36.08pri

Prince, Ellen. 1998. On the limits of syntax, with reference to left-dislocation and topicalization. In *The Limits of Syntax*, Peter Culicover & Louise McNally (eds), 281–302. San Diego CA: Academic Press.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English language*. London: Longman.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Río-Rey, Carmen. 2002. Subject control and coreference in early modern English free adjuncts and absolutes. *English Language and Linguistics* 6(2): 309–323. doi:10.1017/S1360674302000254

Robinson, Ian. 1996. *The Foundations of English Prose: The Art of Prose Writing from the Middle Ages to the Enlightenment*. Cambridge: CUP.

Rohdenburg, Günter. 1995. On the replacement of finite complement clauses by infinitives in English. *English Studies* 76(4): 367–388. doi:10.1080/00138389508598980

Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7: 147–182. doi:10.1515/cogl.1996.7.2.149

Ross, John R. 1967. Constraints on Variables in Syntax. PhD dissertation, MIT. Published as Ross, John R. 1986. *Infinite Syntax!* Norwood NJ: Ablex.]

Runge, Steven E. 2010. *Discourse Grammar of the Greek New Testament*. Peabody MA: Hendrickson.

Shaer, Benjamin, Cook, Philippa, Frey, Werner & Maienborn, Claudia. 2009. Introduction. Dislocation: Concepts, questions, goals. In *Dislocated Elements in Discourse: Syntactic, Semantic, and Pragmatic Perspectives*, Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (eds), 1–27. London: Routledge.

Snider, Neal. 2005. A corpus study of left dislocation and topicalization. Stanford University, Linguistics Department, manuscript. <http://www.stanford.edu/~snider/pubs/qp1.pdf> (30 April 2008).

Tizón-Couto, David. 2012. *Left dislocation in English: A functional-discoursal approach*. Bern: Peter Lang.

Tizón-Couto, David. 2015. A corpus-based account of left-detached items in the recent history of English: Left dislocation vs. left detached-sequences. *English Text Construction* 8(1): 21–64. doi:10.1075/etc.8.1.02tiz

Tizón-Couto, David. 2016. Left-dislocated strings in modern English epistolary prose: A comparison with contemporary spoken left dislocation. In *Outside the Clause: Form and Function of Extra-clausal Constituents* [Studies in Language Companion Series 178] Gunther Kaltenböck, Evelien Keizer & Arne Lohmann (eds), 203–239. Amsterdam: John Benjamins.

Traugott, Elizabeth Closs. 1972. *The History of English Syntax: A Transformational Approach to the History of English Sentence Structure*. New York NY: Holt, Rinehart and Winston.

Traugott, Elizabeth Closs. 2007. Old English left-dislocations: Their structure and information status. *Folia Linguistica* 41(3–4): 405–441.

Traugott, Elizabeth Closs & Trousdale, Graeme. 2013. *Constructionalization and Constructional Changes*. Oxford: OUP. doi:10.1093/acprof:oso/9780199679898.001.0001

van Hoorick, Bart. 1994. Pragmatic positions and the history of English word order. *Working Papers in Functional Grammar* 56. Amsterdam: University of Amsterdam.

van Kemenade, Ans. 1987. *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.

van Kemenade, Ans & Los, Bettelou. 2013. Using historical texts. In *Research Methods in Linguistics*, Devyana Sharma & Rob Podesva (eds), 216–232. Cambridge: CUP.

van Riemsdijk, Henk. 1997. Left dislocation. In Anagnostopoulou, van Riemsdijk & Zwarts (eds), 1–10.

van Riemsdijk, Henk & Zwarts, Frans. 1997. Left dislocation in Dutch and the status of copying rules. In Anagnostopoulou, van Riemsdijk & Zwarts (eds), 13–30.

Wasow, Thomas. 1997. Remarks on grammatical weight. *Language Variation and Change* 9(1): 81–105. doi:10.1017/S0954394500001800

Yaruss, J. Scott. 1999. Utterance length, syntactic complexity, and childhood stuttering. *Journal of Speech, Language, and Hearing Research* 42(2): 329–344. doi:10.1044/jslhr.4202.329

Zeileis, Achim, Meyer, David & Hornik, Kurt. 2007. Residual-based shadings for visualizing (conditional) independence. *Journal of Computational and Graphical Statistics* 16(3): 507–525. doi:10.1198/106186007X237856

Why Scotsmen will drown and shall not be saved

The historical development of *will* and *shall* in Older Scots

Christine Elsweiler

Ludwig-Maximilians Universität München

The divergent use of the modal auxiliaries *will* and *shall* in Scots and Standard English was noticed as early as the eighteenth century. The *Scottish National Dictionary* states that in Scots, first-person *will* instead of *shall* is used for predictions without denoting volition, which, in turn, is said to be often expressed by *shall.* In the second and third person, *will* may express obligation. This study investigates the use of *will* and *shall* in the *Helsinki Corpus of Older Scots* (*HCOS*) with a special focus on the interplay of modal meaning and grammatical person. The corpus data confirm the observations of the *Scottish National Dictionary* for Modern Scots, with a significant drop in first-person volitional *will* in the seventeenth century and fairly high counts for first-person volitional *shall* throughout Older Scots. The prediction uses of both first-person *will* and *shall*, however, remain low throughout the period. In addition, several instances of second- and third-person *will* were found that impose an obligation through an indirect speech act.

Keywords: Older Scots, corpus study, modal auxiliaries, *will*, *shall*, modal meaning, grammatical person

1.   Introduction

Present-day Scots and British Standard English share the core modal auxiliaries *can, must, may, will*, and *shall.* Although, formally, these modals are similar in both varieties, in Scots they are used differently from Standard English (Beal 1997: 365–370; Miller 2008: 304–307) and this also holds true for Scottish Standard English

(Corbett et al. 2003: 2).[1] The differences in the use of the modal auxiliaries in Scots were first observed by 18th-century philosophers such as James Beattie in their lists of Scotticisms, containing items that deviate from the Standard English norm (see Beattie 1797: 29–30). In these lists they remark upon the misapplication of *will* in Scots where Standard English requires *shall*.[2] The divergent use of the two modal auxiliaries in Scots and English forms the basis of the following joke, which plays on the fact that Scots speakers employ the modal auxiliaries *will* and *shall* differently from English speakers:

> A Scotsman was swimming in an English lake. Taken by cramps, he began to sink. He called out for help: "Attention! Attention! I will drown and no one shall save me!" Many people were within earshot, but, being well-brought up Englishmen and women, they honoured his wishes and permitted him to drown.[3]

The joke revolves around the fact that English speakers interpret this *will* in the first person to express volition, whereas Scottish people use first-person *will* for predictions. For English people, the drowning Scotsman is voicing his wish to drown. In the third person, *shall* expresses an obligation for English speakers. In this joke, the passers-by believe that the Scotsman obliges them to refrain from rescuing him. However, by using third-person *shall* the drowning Scotsman merely makes a prediction about his imminent fate. The diverging use of *will* and *shall* in Standard English and Scots can be summed up as follows: According to the prescriptive tradition, in Standard English, *shall* is used in the first person to make predictions. In the second and third person *will* is required. First-person *will*, on the other hand, expresses volition.[4] In Scots, *will* in the first person is used for predictions without denoting volition, which, in turn, is often expressed by *shall.* In the second and third person *will* may express obligation, which in English is usually expressed by *shall* (*Scottish National Dictionary* (*SND*)*,* s.v. *will*).

The history of the modal auxiliaries is well documented for English (e.g. Warner 1993; Fischer 2007: 159–258). Several studies place a special focus on *will* and *shall*

---

1.   The linguistic situation in Scotland is such that there is a linguistic continuum ranging from Broad Scots on one end to Scottish Standard English on the other. Many speakers drift along the continuum between more Scots and more standard variants (Aitken 1984: 522).

2.   Whenever I refer to *will* and *shall* in general in various historical stages, I will use the Present-day English forms *will* and *shall*. Specific instances of *will* and *shall* in the *Helsinki Corpus of Older Scots* (*HCOS*) will be referred to by the spelling variant represented in the corpus texts.

3.   Adapted from <http://www.dailywritingtips.com/the-difference-between-will-and-shall/> (28 February 2015); see also (Sinclair 1782: 73–74).

4.   See also the paper by Ingrid Tieken-Boon van Ostade (1985), where she offers an overview of the development of the rules for the use of *will* and *shall* in 18th-century grammars.

(e.g. Arnovick 1990; Gotti 2003; Nurmi 2003). Gotti et al. (2002), for instance, compares the modal meanings of these two modals in late Middle English and Early Modern English in the *Helsinki Corpus of English Texts* (*HC*). In Gotti's (2006) diachronic study from Old English through to Present-day English, a special focus is placed on the development of the predictive meaning of *will* and *shall*. Gotti finds that in Middle English *shall* was used more frequently than *will* and that in Early Modern English *shall* declined in favor of *will* in the predictive function. In addition to the studies dedicated to the use of *will* and *shall* in various historical stages of English, there are also accounts of their use in other varieties of English where similar distributional patterns of modal meaning and grammatical person hold as in Scots (Beal 1997: 366 and *SND* s.v. *will*). In her comparative study of *will* and *would* in Late Modern Irish and British English, Ronan (2014) finds that *will* is more frequent in Irish English than in British English, particularly in prediction uses. McCafferty and Amador-Moreno's (2012a) quantitative study of first-person *will* and *shall* in an 18th-century letter corpus (*CORIECOR*)[5] of Irish English shows that by the late 18th century *shall* was still preferred over *will*. For Early American English, Kytö (1991: 277–344) demonstrates that first-person *shall* was more frequent in the 17th century, with first-person *will* becoming more popular towards the end of the century. In her brief account of modal meanings, she finds that in British English in the first person, prediction is mainly expressed by *shall* and that *will* is increasingly employed for volitional uses at the expense of *shall*.

For Scots, to date, a description of the historical development of the modal verbs in general and *will* and *shall* in particular is still wanting. Further, although pragmatic uses of *will* and *shall* in the different grammatical persons have been studied (see Arnovick 1990; Gotti 2003) and the question of the applicability of the prescriptive rules for the distribution of *will* and *shall* has been considered (see Fries 1925; Taglicht 1970), there has not been sufficient quantitative research on the diachronic development of modal meaning in conjunction with grammatical person for Standard English nor for other varieties of English.

The present qualitative and quantitative study therefore aims to fill this gap. First, I provide a descriptive account of the semantic development of *will* and *shall* in Older Scots on the basis of corpus data using the *Helsinki Corpus of Older Scots,*[6] 1450–1700 (*HCOS*). Second, I explore to what extent the distribution of *will* and *shall* in the *SND* according to grammatical person and modal meaning features in the *HCOS*. The following research questions form the basis of this study:

---

5. See McCafferty & Amador-Moreno (2012b) for details on the corpus.

6. See <http://www.helsinki.fi/varieng/CoRD/corpora/HCOS/basic.html> and Meurman Solin (1993: 59–124) for a detailed description of the structure and composition of the *HCOS*.

1. Do the *HCOS* data show a trend towards a decline of first-person volitional *will*?
2. Can a concomitant rise of first-person volitional *shall* be observed?
3. Is there a rise in predictive uses of first-person *will*?
4. Is there evidence for the use of *will* in the second and third person to express obligation?

In order to answer these questions, a sample of 1000 instances each of *will* and *shall* from the period 1450–1700 was analyzed to determine which modal meanings are expressed by the two modal auxiliaries. This classification of modal meanings is discussed in the first part of the paper (Section 2). The *HCOS* data and method employed are then outlined (Section 3). Section 4.1 will focus on the development of the modal meanings expressed by *will* and *shall* from 1450 to 1700, and Section 4.2 on the interplay of modal meaning and grammatical person.

2.   Classification of modal meanings

The English modal auxiliaries express a range of modal meanings. Historically, the modal meanings developed from verbs with full referential meaning. The Old English preterite-present verbs **sculan* and *willan* had the lexical meanings 'to owe' and 'to wish', respectively, from which the modal meanings 'obligation' (**sculan*) and 'volition' (*willan*) evolved (Fischer & van der Wurff 2006: 146–152). The use of *shall* and *will* as future tense markers took hold at different stages of Middle English (Visser 1969: Section 1493), with a few attested examples of *shall* already in Old English (see Traugott 1989: 39 and Denison 1993: 298–299). Mustanoja (1960: 490) claims that it is only in later Middle English that *will* is employed for the expression of future tense without volitional overtones.

For the classification of modal meanings many linguists draw a basic binary distinction between root modality on the one hand and epistemic modality on the other. Root modality comprises meanings such as obligation, volition, and propensity. Epistemic modality, by contrast, is defined as an assessment of the truth value of a proposition (e.g. Coates 1983: 18–22; Leech 2004: 84). The root meanings are taken to be more basic than the epistemic meanings, which are believed to have developed out of the root meanings diachronically (Traugott 1989: 36–37). Palmer (2001) adopts a division between epistemic modality and what he terms "event" modality, which subsumes dynamic and deontic modality. The distinction between the latter two categories is based on whether the source for the modality is internal to a discourse-participant, as is the case with dynamic meanings such as propensity and volition, or whether it is external, as pertains to deontic meanings such as obligation. Another approach (e.g. Huddleston & Pullum 2002; Collins 2009)

to the classification of modal meanings is a tripartite division between epistemic, deontic, and dynamic modality. This distinction forms the basis of the classification of the modal meanings expressed by the two modal auxiliaries *will* and *shall* in the present study. Below, I list the modal meanings relevant to this investigation of *will* and *shall* with illustrative examples taken from the *HCOS*.

Epistemic modality comprises the modal meanings 'prediction' (1) and 'probability judgment' (2).

(1)   *I trast ze **sall heyre** quhow this promotone now imprecat be zonn dyssatfull Byschep of Morray sall turne to our weyll.*
'I trust that you will hear how this promotion now invoked by that deceitful Bishop of Murray will turn in our favor.'

(SC1, Official Correspondence, Gavin Douglas to Adam Williamson, 1515, <P 71>)

Predictions can be identified by means of the modal paraphrase *predict that* and by the presence of other indicators such as stance verbs in the matrix clause, e.g. *trast* 'trust' in (1).

Probability judgments normally make a statement about the present. A possible paraphrase would be *believe that* or a modal adverb such as *doutles* 'doubtless'.

(2)   *and because it **will be** very deare living in Paris, ye may stay at Orleance w[{nt{] ill the begining of December.*
'and because it will be very expensive living in Paris, you may stay at Orleans until the beginning of December.'

(SC3, Private Correspondence, Sir Thomas Steuart of Grandtully to his son, 1669, <P 192>)

Obligation, which typically derives from an external authority, is classified as deontic (3).

(3)   *ITEM it is statut & ordanit in þis present parlyament þat na man **sal weir** silkis in tyme cumyng in gown doublate & clokis except knychtis menstrallis & herraldis.*[7]
'Further, it is decreed and ordained in this present parliament that no man shall wear silk in time to come in gown, doublet and cloak, except knights, minstrels and heralds.'   (SC0, Parliament, *The Acts of the Parliaments of Scotland*, Edinburgh, 1471, <P 100.C1>)

The presence of verbs such as *statut and ordanit* 'decreed and ordained' in the matrix clause help identify this use of *shall*.

---

7.   The *HCOS* coding conventions for abbreviations were replaced by the full spelling in this and the following examples.

Dynamic modality as defined by Huddleston & Pullum comprises "dispositions or properties of the subject-referent" (2002: 192). It covers the meanings volition (4) and propensity (5).

(4) *My Lord, your ladyship's brother, is weill, and **will sie** yow this nixt yeir.*
'My lord, your ladyship's brother, is well and will see you this coming year.'
(SC2, Private Correspondence, Sir John Grant
of Freuchie to Lilias Murray, 1631, <P 55>)

Volition always originates from an animate subject. The volitive use of *will* in (4) may be glossed by 'intend to'. The volitional meanings expressed by the two modal auxiliaries can be further sub-categorized into the more specific meanings 'intention' for both *will* and *shall*, and 'willingness' and 'refusal' (i.e. negated willingness) for *will* only. In the quantitative analysis of the *HCOS* material in Section 4, all these sub-categorizations will be subsumed under 'volition'.

Propensity refers to "characteristic or habitual behaviour of animates" or "general properties of inanimates" (Huddleston & Pullum 2002: 194)*,* as in (5).

(5) *And thar'fore to ken gude wateris thir ar the takenis /. the gude water' js jn gude place tane … quhite and clere /wele gustit and licht / and sone **will be** hate and sone cule agayne to the propre kynde.*
'And therefore, to know good waters, those are the signs. The good water is taken in a good place, … white and clear, of good taste and light; and soon it will be hot and soon cool again to the proper kind.'
(SC0, Education, *The Buke of the Gouvernaunce of Princis*, 1456, <P 106>)

It is a timeless characteristic of the water described in (5) that it turns hot and then cold again. In the *HCOS*, however, the instances of *will* and *shall* expressing characteristics and habits are rare, so that the propensity meanings are subsumed under "other" in the empirical analysis of the data in Section 4.

Beside the modal meanings introduced so far, which also apply to Present-day English usage, I have encountered a further use, described in Visser (1969: Sections 1516, 1519), where the modal auxiliary is typically employed in subordinate clauses, adding little modal coloring.

(6) *And this foirsaid pane to be applyit to the pure folkis be thame that **salbe depute** collectouris thairof.*
'And this aforementioned penalty to be applied to the poor folks by them that will be assigned collectors thereof.'
(SC1, Parliament, *The Acts of the Parliaments of Scotland*, 1551, <P 485.C2>)

In (6), *salbe depute* in the relative clause modifying *thame* signifies that the collectors are as yet undetermined. The relative clause contains a statement about a future

event, but the modal construction can neither be interpreted as making a prediction nor expressing volition. Since the focus is on the indication of future time, I label this use low-degree modality (cf. Huddleston & Pullum 2002: 180).[8] The modality is not zero, however, since future events are not entirely certain to happen. Visser (1969: Section 1516) notes that in unambiguous cases Present-day English uses the present tense instead of *shall*.

   Apart from the categories introduced above, I included indeterminate categories in my analysis for cases which can be interpreted as expressing more than one modal meaning (cf. Coates 1983: 16–17). Example (7), for instance, can be read as either intention on the part of the messenger or as a prediction Bothwell makes to the Queen Dowager.

(7)   *And quhat skayth and displesour I have sustenit presently this berar **will schaw** unto your grace at lenht.*
      'And what harm and displeasure I have suffered, this messenger will show to your grace at length in person.'
                                          (SC1, Official Correspondence, Earl Bothwell to the
                                                       Queen Dowager, before 1556, <P 440>)

Although the phrasing *this berar will schaw … at lentht* is formulaic in character, referring to the fact that the messenger will ensure that the message stays confidential and therefore favoring a prediction reading, the interpretation of *will* as denoting intention is also possible. In the present study, this indeterminate use is labeled 'prediction/intention'. A further case of indeterminacy obtains in examples of indirect speech acts with *will* where a prediction is so binding that it can also be interpreted as an obligation, as in (8) (see Huddleston & Pullum 2002: 194).

(8)   *Ye **will not faill** to send hame ane atturnay to tak sesing for it.*
      'You will not fail to send home an attorney to take possession for it.'
                                          (SC2, Private Correspondence, Lady Barnbarroch
                                                       to her husband, 1587, <P 394>)

These indirect speech acts fall into the category 'prediction/obligation', which further contains examples of *shall* that may be interpreted to express both prediction and obligation.

---

8.   Note that my interpretation of low-degree modality expands Huddleston & Pullum's definition by not only including the degree of the modal meaning but also the kind of modality.

3.   Data and method

The present study was conducted using the *Helsinki Corpus of Older Scots* (*HCOS*), a diachronic corpus comprising 834,200 words, which spans the time period from 1450 to 1700. The *HCOS* is further divided into four sub-periods SC0 (1450–1500), SC1 (1500–1570), SC2 (1570–1640), and SC3 (1640–1700). Whereas the texts included in the first sub-period SC0 only amount to a total of 85,100 words, the word counts for the remaining three sub-periods are 201,800 (SC1), 305,900 (SC2), and 241,400 (SC3). For each of these periods, the sub-corpus includes a range of up to fourteen different genres. Ideally, all four sub-periods of the *HCOS* would have exactly the same range of genres and texts. However, only SC2 and SC3 cover all fourteen genres, with the first two sub-periods having a more limited range. Further, the individual texts featured in the *HCOS* are of varying length. For my investigation of the use of *will* and *shall*, I conducted a regular expression search for both modal auxiliaries based on the spelling variants given in the *Dictionary of the Older Scottish Tongue* (*DOST*, s.vv. *will* and *sall*).[9] Then I drew a random stratified sample of 1000 instances of *will* and *shall* each, i.e. the data points were randomly drawn in such a way that their frequency in the sample reflects the ratio of their overall frequency in the corpus across all the texts in the different time periods.[10] This was important to avoid sampling bias. The cliticized form *'ll*, which occurs twenty-three times, cannot clearly be attributed to either *will* or *shall* and was therefore excluded from the analysis. In a next step, the instances of the two modals were classified according to modal meaning by applying the criteria described in Section 2.

4.   *Will* and *shall* in the *Helsinki Corpus of Older Scots*

The following discussion will first describe the diachronic development of the modal meanings of *will* and *shall* in the *HCOS* sample. In a second step, I will undertake an analysis of the patterns arising from the conjunction of grammatical person with modal meaning. A cross-genre survey will be presented of the uses of *will* and *shall* in the different grammatical persons. I will, however, single out distributional preferences in particular genres and specific usage patterns, making use of illustrative examples.

---

9.   It was expected that by using the spellings given in *DOST* in the regular expression search, the vast majority of spelling variants for *will* and *shall* would be covered except for possibly a few idiosyncratic spellings in genres such as, for instance, correspondence.

10.   The number of instances of *will* and *shall* actually exceeds the target number of 1000 in both cases, since all occurrences of the two modal auxiliaries in a corpus text were kept if the sampled output was less than ten.

4.1 The modal meanings expressed by *will* and *shall* in the *HCOS*

Table 1 below shows the distribution of modal meanings of *will* in the four sub-periods of Older Scots. The table includes the raw numbers, as retrieved from the *HCOS* sample, in the left column and the percentage share of the respective modal meanings in the right column. In the first two sub-periods SC0 (1450–1500) and SC1 (1500–1570), *will* predominantly expresses volition with a share of 52.63% (SC0) and 45.49% (SC1) of all uses. There is a marked decrease in SC2 (1570–1640) and SC3 (1640–1700), where volitional *will* drops to 29.33% in SC2 and eventually 21.33% in SC3. The prediction uses of *will*, on the other hand, only have a share of 15.79% in SC0 and then rise to 47.73 and 43.71% in SC2 and SC3, respectively. In order to test the statistical significance of the drop of volitional *will* and the rise of the predictive uses, the four-way period division was reduced to a two-way division with SC0 and SC1 representing the earlier period and SC2 and SC3 the later period. A Pearson's Chi-squared test with Yates' continuity correction showed that there is a highly significant association between the time period and whether or not volition is present $X^2 (1) = 44.8184$, $p < 0.001$. The same is true of the increase of prediction uses, where the test revealed a highly significant association between the time period and the presence of prediction $X^2 (1) = 32.5218$, $p < 0.001$.[11]

Table 1. Changing distribution of modal meanings of *will* by time period
(n = raw frequency)

	volition		prediction		probability		volition/ prediction		prediction/ obligation		low modality		other		Total
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n
SC0	40	52.63	12	15.79	1	1.32	0	0	4	5.26	7	9.21	12	15.79	76
SC1	121	45.49	81	30.45	9	3.38	23	8.65	10	3.76	11	4.14	11	4.14	266
SC2	110	29.33	179	47.73	26	6.93	22	5.87	20	5.33	11	2.93	7	1.87	375
SC3	61	21.33	125	43.71	36	12.59	24	8.39	2	0.7	11	3.85	27	9.44	286

Figure 1, representing the main modal meanings encoded by *will*, shows that the turnaround between predominantly volitional uses and predominantly predictive uses happens between SC1 and SC2, the sub-periods between 1500 and 1640. The drop of volitional uses and concomitant rise of predictive uses of *will* in Older Scots confirms previous findings for English, which showed an increase of predictive *will* between Late Middle English and Early Modern English (Gotti 2006: 101–102, 106, 112).

---

11. See Table 9 in the Appendix for all the Pearson's Chi-squared test results for the changing distributions of modal meanings.
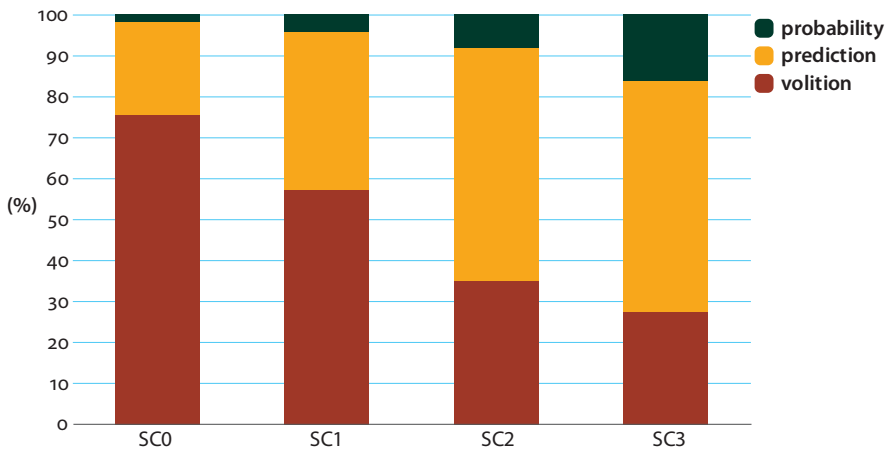
Figure 1. *Will* – distribution of main modal categories by time period (1450–1700)

A comparison of the development of the predictive uses of *will* and *shall* in the *HCOS* data (see Figure 2) illustrates that it was also between SC1 and SC2 that predictive *will* became more frequent than predictive *shall*.
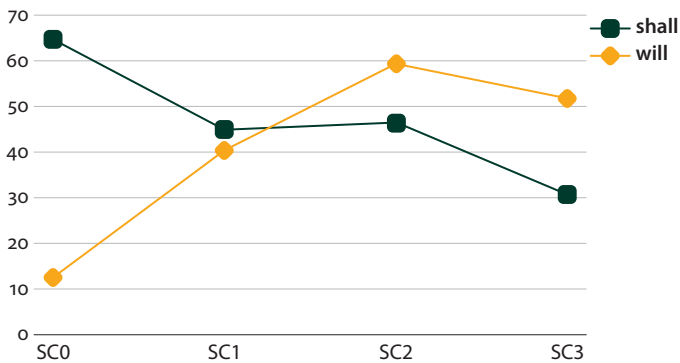
Figure 2. The diachronic development of predictive *will* and *shall* (normalized frequencies per 100,000 words)

The other interesting development in the epistemic category is the gradual increase of the relative share of *will* for probability judgments from just 1.32% in the 15th century (SC0) to 12.59% in the 17th century (SC3) (see Table 1). A comparative glance at the distribution for *shall* (see Table 2 and Figure 3) reveals that *will* is indeed the preferred modal for probability judgments. The uses for *shall* do not exceed the 2%- mark in any of the sub-periods. *Will*, for its part, is as rare as *shall* in the first sub-period SC0 but then establishes itself as a probability marker in the course of the 17th century. This

rise between the time periods SC0/SC1 on the one hand and SC2/SC3 on the other also proves statistically highly significant X^2 (1) = 13.1452, $p < 0.001$.

The distribution of *shall* represented in Table 2 shows that the obligation meaning is predominant in all four sub-periods. However, this use sees a decline from nearly half of all the uses in the 15th century (SC0) to some 30% in the 17th century (SC2 and SC3), which is statistically highly significant X^2 (1) = 19.8988, $p < 0.001$. A similar drop in obligation uses has been reported for *shall* in Late Middle English (Gotti 2006: 101, 112).

Table 2. Changing distribution of modal meanings of *shall* by time period (n = raw frequency)

	obligation		volition		prediction		probabity		prediction volition/		obligation prediction/		low modality		other		Total
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n
SC0	109	49.77	8	3.65	62	28.31	4	1.83	0	0.00	14	6.39	15	6.85	7	3.20	219
SC1	146	43.07	44	12.98	90	26.55	4	1.18	9	2.65	3	0.88	33	9.73	10	2.95	339
SC2	145	35.02	29	7.00	139	33.57	3	0.72	9	2.17	6	1.45	65	15.70	18	4.35	414
SC3	112	31.64	42	11.86	74	20.90	4	1.13	6	1.69	18	5.08	91	25.71	7	1.98	354

Predictive *shall* remains stable all through the sub-periods with a share of roughly one third, with only slight random variation. The share of volitional *shall* (cf. Section 4.2.1) remains at a relatively low level in all four periods with no significant change. The rise of instances of *shall* having a low degree of modality from 6.85% in the 15th century (SC0) to 25.71% in the 17th century (SC3), however, is an interesting change. This increase likewise proves statistically highly significant X^2 (1) = 33.1516, p <0.001.
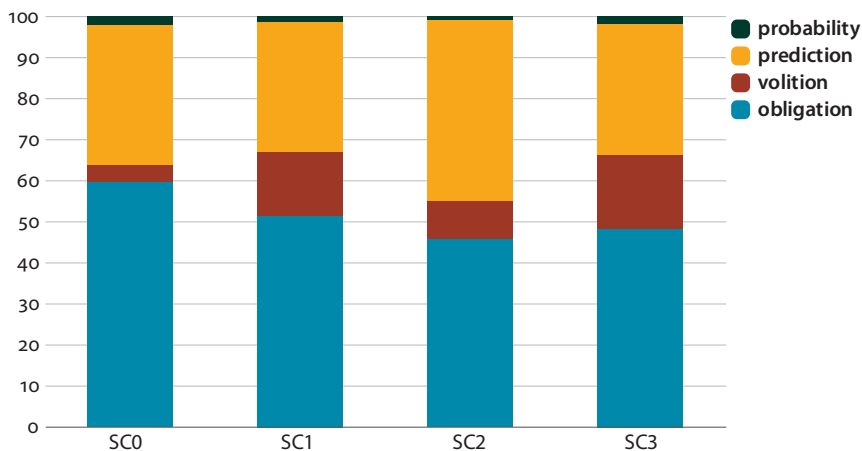
Figure 3. *Shall* – distribution of main modal categories by time period (1450–1700)

4.2 The distribution of *will* and *shall* by grammatical person

This section will discuss the trends in the *HCOS* data when considering the interplay of modal meaning and grammatical person. In particular, the research questions formulated at the beginning of this paper will be addressed.

4.2.1 *First person*[12]

Based on the claims in the *SND* definition s.v. *will*, the hypothesis is that first-person volitional *will* declines in the course of the Older Scots period and that first-person volitional *shall* increases. Accordingly, a rise is expected for first-person predictive *will* during the periods covered by the *HCOS* and a drop for first-person predictive *shall*.

Volition
Table 3 shows the development of first-person volitional *will* and *shall*. The left column states the raw frequencies of *will* and *shall* in the *HCOS* sample. The percentages in the right column report the share of the volitional meanings relative to the other modal meanings, i.e. the total of all first-person instances of *will* and *shall*, in the relevant sub-period. The percentage shares indicate that the use of first-person volitional *will* declines in the course of the Older Scots period. There is a drop from 86.54% in SC1 (1500–1570) to 63.16% in SC3, which is statistically significant X^2 (1) = 5.3364, $p < 0.05$. First-person volitional *shall*, on the other hand, remains relatively stable with fairly high counts throughout all four periods.

Table 3. The changing distribution of first-person volitional *will* and *shall* (n = raw frequency; % = percentage share relative to all modal meanings)

	1P volitional *will*		1P volitional *shall*	
	n	%	n	%
SC0	4	100	7	58.33
SC1	45	86.54	37	69.81
SC2	58	74.68	26	41.94
SC3	24	63.16	39	73.58

Figure 4, which does not consider the first sub-period SC0 since there are very few instances of either *will* or *shall*, illustrates that at the end of the 17th century first-person volitional *will* was used less frequently than *shall*.

---

12. See Table 10 in the Appendix for all the Pearson's Chi-squared test results for the changing distributions of modal meanings of first-person *will* and *shall*.
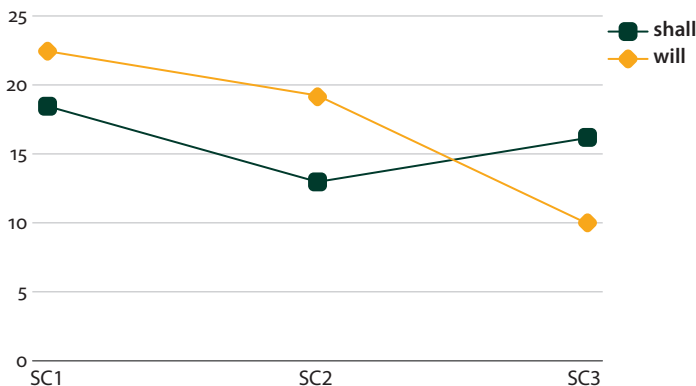
Figure 4. The development of first-person volitional *will* and *shall* from SC1 to SC3 (normalized frequencies per 100,000 words)

These findings corroborate my hypothesis that there is a drop in volitional *will* in the first person. For first-person volitional *shall*, however, there is no significant change.

Both first-person volitional *will* and *shall* are mainly found in letters as well as in pamphlets, which may be explained by the nature of these genres. First-person volitional *will* predominantly expresses intention, as in (9).

> (9)   *And the letteris is sa generall that I knawe nocht quhair wyth aw he charges me,*
>   *and in respect of the generalitie I **will get thaim suspendit** vpoun cautioun.*
>   'And the letters are so general that I know not what he charges me with. And
>   with respect to their generality, I will get them suspended upon caution.'
>   (SC2, Private Correspondence, Sir William Scott
>   of Balwearie to John Wemyss, 1569, <P 68>)

With intentional *will*, the focus is on planned future events. In (9), the letter writer voices his plan to delay the payments imposed on him in the letters he refers to. Beside the intention use, first-person *will* is also employed for the expression of willingness (10) and refusal (11), which focus on the current state of mind of the discourse-participant (Collins 2009: 132). The modal paraphrase for *will* in (10) is *am willing to* and in (11) *I refuse to.*

> (10)   *Now to mak ansuer wnto the first pairt of your letter desyring me to haue yow*
>   *excused that ye did not go to court efter the resett of my letter. Treulie (brother)*
>   *I **will tak** your excuis in verie good pairt,*
>   'Now to answer to the first part of your letter desiring me to have you excused
>   that you did not go to court after the receipt of my letter. Truly (brother) I will
>   take your excuse in good spirits,'   (SC2, Private Correspondence, John,
>   Twelfth Earl of Sutherland, to his brother, 1615, <P 114>)

(11)   *For tho' I entend to give her noe more fruitless advice, yitt I **will not consent** to let her marry anie body that apears to be ane enemy both to the Government of Church and State.*

'For although I intend to give her no more fruitless advice, yet, I will not consent to let her marry anybody that appears to be an enemy both to the government of church and state.'   (SC3, Private Correspondence, Margaret, Countess of Wemyss, to Anne, Countess of Leven, 1696, <P 157>)

First-person volitional *shall* is exclusively employed to express intention but never for the expression of willingness or refusal. Two particular intention uses stand out. One is that first-person *shall* is preferred over *will* for commissive speech acts, i.e. promises and pledges. Given the nature of these commissive speech acts, which involve an element of obligation, it is not surprising that they occur almost exclusively in letters, where the letter writer undertakes an obligation towards the addressee. They are particularly prominent in the official letters in the SC1 sub-period (1500–1570) with fourteen instances. An example is (12), with Earl Bothwell making a pledge to carry out the Queen Dowager's orders to the best of his ability.

(12)   *And sua I have schawin to the berar my mynd in your grace commandmentis quhilkis I **sall fulfyll** at my hichtast power.*

'And so I have shown to this messenger my intention regarding your Grace's commandments, which I will fulfill to the best of my ability.'
(SC1, Official Correspondence,
Earl Bothwell to the Queen Dowager, 1544, <P 115>)

Second, there are 29 instances of volitional *shall* in the first person where it is employed as a device for discourse organization (13), in particular in pamphlets and sermons.

(13)   *The other extremity that commes against the Apostles order is pompe. … But for the more particular deciphering thereof, we **shall range** it in two rankes: the one sort of it being ciuile or seculare, the other superstitious.*

'The other extremity that comes against the Apostles' order is pomp. … But for a more detailed description of it, we will arrange it in two kinds: the one sort of it being civil or secular, the other superstitious.
(SC2, Pamphlet, William Birnie, *The Blame of Kirk-Buriall*, 1606, <P 11>)

The author of the pamphlet uses *shall* to inform the reader about the structure of the upcoming passage. First-person *will* is an alternative to *shall* in this use. There is, however, an observable difference in the distribution of *will* and *shall* in the different time periods. First-person *will* is the preferred device in SC1 (1500–1570) with 7 instances against 2 for *shall*. In SC3 (1640–1700), on the other hand, there is a preference for *shall* with 16 instances against 5 for *will*, which confirms the trend displayed in Figure 4.

Prediction

As can be seen from Table 4, the numbers of both first-person predictive *will* and *shall* remain low throughout the Older Scots period and in all the sub-periods. It is only in SC2 that the percentage share of *shall* is relatively high (32.26%) (twice as high as the second highest percentage share of 16.67% for *shall* in SC0). In the first three sub-periods SC0 to SC2 (1450–1640), first-person *shall* is more frequently used for predictions than *will*.

Table 4. The changing distribution of first-person predictive *will* and *shall* (n = raw frequency; % = percentage share relative to all modal meanings)

	1P predictive *will*		1P predictive *shall*	
	n	%	n	%
SC0	0	0	2	16.67
SC1	2	3.85	8	15.09
SC2	11	12.66	20	32.26
SC3	5	13.16	7	13.21

A closer look at the predictive use of *will* and *shall* in the different genres reveals that *shall* is the preferred variant in sermons, as in (14).

(14)   *He will be with zow, and blys all the warkis of zour handis, as he hes promysit in his worde, quhairof we **sall haue** better occasioun to speik in the thrid heid of our Text.*
'He will be with you, and bless all the work of your hands, as he has promised in his word, whereof we shall have more opportunity to speak in the third head of our text.'   (SC2, Sermon, Tracts by David Fergusson, Minister of Dunfermline, 1571, <P 77>)

In letters, both predictive *will* and *shall* are found without any noticeable preferences.

4.2.2   Second person

Prediction/obligation

The dictionary entry for *will* in the *SND* states that in the second and third person, Scots *will* can be employed to denote obligation, i.e. it is used in contexts where in English *shall* would be employed. The only attestations of this use in the *HCOS* data belong to the indeterminate category 'prediction/obligation'. As was argued in Section 2 above, predictions such as in (15) can be interpreted as an indirect speech act imposing an obligation upon the addressee if the author has sufficient authority over them.

(15)   *Ye **will not faill** to expeid this to the effeck the sesing may be tane before witsonday*
       'You will not fail to complete this to the effect that possession may be taken
       before Whitsunday.'   (SC2, Private Correspondence, Lady Barnbarroch
                                                    to her husband, 1587, <P 394>)

These uses of second-person *will* are not particularly numerous in the corpus with altogether 13 instances, most of which can be found in letters in SC2 (1570–1640). The most popular collocation is *ye will not faill* as in (16), which, however, is exclusively used by one author.

A comparison of predictive *will* and *shall* in the second person reveals a preference for *shall* in all sub-periods except for SC2 (see Table 5), where predictive *will* has a percentage share of 67.12% of all instances of second-person *will*.

Table 5. The changing distribution of second-person predictive *will* and *shall*
(n = raw frequency; % = percentage share relative to all modal meanings)

	2P predictive *will*		2P predictive *shall*	
	n	**%**	**n**	**%**
SC0	0	0	9	52.94
SC1	12	38.71	19	35.19
SC2	49	67.12	38	54.29
SC3	15	42.86	23	69.70

24 of the 49 instances of predictive *will* in SC2 feature in the genre "handbook", where there is an apparent stylistic preference for *will* rather than *shall*; an example is (16).

(16)   *You shall ever seeke the breadth vppon the head of the Table, and the length vppon
       the left side of that Columne, and in the broade roome against the length, you
       will finde the aunswer what the compt extendes to.*
       'You shall always seek the breadth on the head of the table, and the length on
       the left side of that column, and in the broad space facing the length you will
       find the answer indicating what the count extends to.'
                    (SC2, Handbook, Huntar, *Weights and Measures*, 1624, <P 14>)

The rule in (16) is formulated by employing *shall* in its obligation meaning and then the predicted consequence is expressed by *will*.

4.2.3 *Third person*[13]

Prediction/obligation
The uses of *will* in the category prediction/obligation in the third person are similar to those in the second person, as is illustrated in (17).

> (17) *And ordanis þame to proceid and do Justice þerintill vnto þe finall end and decisioun þerof As þai* **will ansuer** *to god..*
> 'And ordains them to proceed and do justice thereto until the final end and decision thereof, as they will answer to god.'
> (SC2, Parliament, *The Acts of the Parliaments of Scotland*, 1587–1621, <P 498.C2>)

In the third person, these indirect speech acts are particularly frequent in legal texts and official letters from SC1 to SC3 (1500–1700), with the most common formula being *as they will ansuer to God*.

Prediction
The highest number of predictive uses of *will* and *shall* occurs in the third person, as becomes evident when comparing the raw frequencies in Tables 4 and 5 above and in Table 6 below. As can furthermore be observed, the increase in the percentage share of prediction uses of third-person *will* (relative to all modal meanings) between SC0/SC1 and SC2/SC3 is statistically highly significant $X^2 (1) = 24.0161$, $p < 0.001$.

Table 6. The changing distribution of third-person predictive *will* and *shall* (n = raw frequency; % = percentage share relative to all modal meanings)

	3P predictive *will*		3P predictive *shall*	
	n	**%**	**n**	**%**
SC0	12	18.18	51	26.84
SC1	67	36.61	63	27.16
SC2	119	53.36	81	28.72
SC3	105	49.3	44	16.42

For third-person predictive *shall*, on the other hand, no significant change can be observed.

---

13. See Table 11 in the Appendix for all the Pearson's Chi-squared test results for the changing distributions of modal meanings of third-person *will* and *shall*.

Volition

The drop in volitional uses of *will* in the course of the Older Scots period (cf. Section 4.1) is not restricted to the first person but also highly significant with the third person X^2 (1) = 47.9807, $p < 0.001$.

Probability judgments

The increase in the use of *will* for probability judgments observed in Section 4.1 happens predominantly in the third person. The rise of third-person *will* in this use between the first two sub-periods and the last two (see Table 7), from 1.52% (SC0) and 4.37% (SC1) to a percentage share of 11.21% and 15.49% in SC2 and SC3, respectively, is highly significant X^2 (1) = 15.778, $p < 0.001$.

Table 7. The changing distribution of third-person *will* and *shall* used for probability judgments (n = raw frequency; % = percentage share relative to all modal meanings)

	3P probability *will*		3P probability *shall*	
	n	**%**	**n**	**%**
SC0	1	1.52	2	1.05
SC1	8	4.37	3	1.29
SC2	25	11.21	2	0.71
SC3	33	15.49	2	0.75

This use of third-person *will* occurs across several genres, but is especially prominent in pamphlets and letters, as in the letter-closing lines in (18).

(18)   *Ye may may* [sic] *send yowr letters to Aldie so long as he is at London, which* **will be** *the surest way.*
'You may send your letters to Aldie as long as he is in London, which will be the surest way.'   (SC3, Private Correspondence, Sir Thomas Steuart of Grandtully to his son, 1669, <P 184>)

The letter writer is making an assumption about which method is best to get his son's letters to Edinburgh. *Will* could be replaced by the modal adverb *probably* to mark the relative clause as a probability judgment.

Low-degree modality

Finally, I will discuss the rise in the use of *shall* with little modal coloring in the course of the Older Scots period. Table 8 demonstrates that *shall* expressing only a low degree of modality rises from a share of 7.37% of all modal meanings in SC0 (1450–1500) to a share of 33.96% in SC3 (1640–1700).

Table 8. The changing distribution of third-person *shall* denoting chance (n = raw frequency; % = percentage share relative to all modal meanings)

	3ᴘ low modality *shall*	
	n	**%**
SC0	14	7.37
SC1	31	13.36
SC2	61	21.63
SC3	91	33.96

This use is especially prominent in legal texts, where it exclusively occurs in subordinate clauses and among these most commonly in conditional (19) and in relative clauses (20).

(19) *And **giff It salhappin** anye man to win anye sowmes of money at Carding Or dyceing Attoure the soume of ane hundereth merkis within the space of Tuentie four houres.*
'And if it was to happen to any man to win any sums of money at card or dice games around the sum of one hundred marks within the space of twenty-four hours.' (SC2, Parliament, *The Acts of the Parliaments of Scotland*, 1587–1621, <P 613.C2>)

(20) *AND the saidis strangers and workmen presentlie within this realme or that **salhappin** to cum within the same.*
'And the aforementioned strangers and workmen currently in this realm or who will happen to come into the same.' (SC2, Parliament, *The Acts of the Parliaments of Scotland*, 1587–1621, <P 509.C1>)

Examples (19) and (20) show that *shall* often collocates with the verb *happin*. Both in (19) and (20) *shall* may be interpreted as a redundant means to reinforce the conditionality expressed by the conjunction *gif* in the *if*-clause and the verb *happen* in both clauses.

5. Conclusion

By way of conclusion, I will firstly address the four research questions forming the basis of my analysis and secondly compare my findings for Older Scots to findings for other varieties of English.

The present study of the development of modal meanings in the different grammatical persons confirm the trends described in the *SND* s.v. *will*. Across the Older Scots period, there is a significant decrease in volitional *will*, compared with its

other meanings; this fall is evident both in the first person and in the third person. My first research question (Do the *HCOS* data show a decline of first-person volitional *will*?) can therefore be answered affirmatively. At the same time, first-person *shall*, while being relatively frequent, does not show a concomitant rise (second research question). Actually, *shall* shows no clear diachronic trend, neither generally nor in the first person. With regard to first-person predictive *will*, no increase can be observed (third research question). The overall increase in predictive *will*, which is noticeable across all genres, occurs in the third-person uses. Similarly, third-person *will* also shows a rise in probability judgments. There are no marked changes in predictive *shall* generally (with a percentage share of roughly one third relative to the other modal meanings). The final research question sought to establish if second- and third-person *will* is used to express obligation. Several instances of second- and third-person *will* in the indeterminate category 'prediction/obligation' were found, which represent indirect speech acts imposing an obligation upon the interlocutor.

My findings for Older Scots regarding the general drop in frequency of volitional *will* and the rise of instances of predictive *will* match the development of the modal meanings of *will* in Early Modern English (see Gotti 2006). Interestingly, Ronan's (2014) study of Late Modern British English, based on four genres represented in the *ARCHER* corpus, does not corroborate Gotti's results since her data show that volition is the predominant modal meaning of *will* with 49 per cent of all examples, followed by prediction with 33 per cent. Her findings for Late Modern Irish English, however, confirm the pattern I found for Older Scots, with prediction accounting for 52 per cent of all instances and volition for 28 per cent (see Ronan 2014: 247).

For *shall*, the obligation meaning is predominant in all four sub-periods. Still, a decline in this use of *shall* can be observed. A similar drop of the frequency of *shall* expressing obligation was observed by Gotti (2006) for Early Modern English. Another interesting development for *shall* is the rise in the seventeenth century of its uses expressing a low degree of modality. In many of the cases in this category, neither *will* nor *shall* would be employed in Present-day English.

References

Aitken, Adam J. 1984. Scots and English in Scotland. In *Language in the British Isles*, Peter Trudgill (ed.), 517–532. Cambridge: CUP.

Arnovick, Leslie K. 1990. *The Development of Future Constructions in English: The Pragmatics of Modal and Temporal* will *and* shall *in Middle English*. Bern: Peter Lang.

Beal, Joan. 1997. Syntax and morphology. In *The Edinburgh History of the Scots Language*, Charles Jones (ed.), 335–377. Edinburgh: EUP.

Beattie, James. 1797. *Scoticisms, Arranged in Alphabetical Order, Designed to Correct Improprieties of Speech and Writing*. Edinburgh: Printed for the booksellers.

Coates, Jennifer. 1983. *The Semantics of the Modal Auxiliaries*. London: Croom Helm.

Collins, Peter. 2009. *Modals and Quasi-modals in English*. Amsterdam: Rodopi. doi:10.1163/9789042029095

Corbett, John, McClure, J. Derrick & Stuart-Smith, Jane (eds). 2003. *The Edinburgh Companion to Scots*. Edinburgh: EUP.

Denison, David. 1993. *English Historical Syntax: Verbal Constructions*. London: Longman.

DOST = Craigie, William, Aitken, A. J., Stevenson, James A. C. & Dareau, Marace (eds). 1932. *A Dictionary of the Older Scottish Tongue*. <http://dsl.ac.uk>

Fischer, Olga. 2007. *Morphosyntactic Change: Functional and Formal Perspectives*. Oxford: OUP.

Fischer, Olga & van der Wurff, Wim. 2006. Syntax. In *A History of the English Language*, David Denison & Richard Hogg (eds), 109–198. Cambridge: CUP. doi:10.1017/CBO9780511791154.004

Fries, Charles C. 1925. The periphrastic future with *shall* and *will* in Modern English. *PMLA: Publications of the Modern Language Association of America* 40(4): 963–1024. doi:10.2307/457534

Gotti, Maurizio, Dossena, Marina, Dury, Richard, Facchinetti, Roberta & Lima, Maria (eds). 2002. *Variation in Central Modals: A Repertoire of Forms and Types of Usage in Middle English and Early Modern English*. Bern: Peter Lang.

Gotti, Maurizio. 2003. Pragmatic uses of *shall* and *will* for future time reference in Early Modern English. In *English Modality in Context: Diachronic Perspectives*, David Hart (ed.), 109–170. Bern: Peter Lang.

Gotti, Maurizio. 2006. Prediction with SHALL and WILL: A diachronic perspective. In *The Changing Face of Corpus Linguistics*, Antoinette Renouf & Andrew Kehoe (eds), 99–116. Amsterdam: Rodopi.

HCOS = Meurman-Solin, Anneli. 1995. *The Helsinki Corpus of Older Scots*. Helsinki: Department of Modern Languages, University of Helsinki.

Huddleston, Rodney & Pullum, Geoffrey. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.

Kytö, Merja. 1991. *Variation and Diachrony, with Early American English in Focus*. Frankfurt: Peter Lang.

Leech, Geoffrey. 2004. *Meaning and the English Verb*, 3rd edn. Harlow: Pearson Education.

McCafferty, Kevin & Carolina Amador-Moreno. 2012a. 'I will be expecting a letter from you before this reaches you': A corpus-based study of *shall/will* variation in Irish English correspondence. In *Letter Writing in Late Modern Europe* [Pragmatics & Beyond New Series 218], Marina Dossena & Gabriella Del Lungo Camiciotti (eds), 179–204. Amsterdam: John Benjamins. doi:10.1075/pbns.218.11mcc

McCafferty, Kevin & Amador-Moreno, Carolina. 2012b. A corpus of Irish English Correspondence (*CORIECOR*): A tool for studying the history and evolution of Irish English. In *New Perspectives on Irish English*, Bettina Migge & Máire Ní Chiosáin (eds), 265–287. Amsterdam: John Benjamins. doi:10.1075/veaw.g44.13mcc

Meurman-Solin, Anneli. 1993. *Variation and Change in Early Scottish Prose: Studies Based on the Helsinki Corpus of Older Scots*. Helsinki: Suomalainen Tiedeakatemia.

Miller, Jim. 2008. Scottish English: Morphology and syntax. In *Varieties of English: The British Isles*, Bernd Kortmann & Clive Upton (eds.), 299–327. Berlin: Mouton de Gruyter.

Mustanoja, Tauno. 1960. *A Middle English Syntax. Part I*: Parts of speech. Helsinki: Société Néophilologique.

Nurmi, Arja. 2003. *Youe shall see I will conclude in it*: Sociolinguistic variation of WILL/WOULD and SHALL/SHOULD in the sixteenth century. In *English Modality in Context: Diachronic Perspectives*, David Hart (ed.) 89–108. Bern: Peter Lang.

Palmer, Frank. 2001. *Mood and Modality*. Cambridge: CUP. doi:10.1017/CBO9781139167178

Ronan, Patricia. 2014. Tracing uses of *will* and *would* in Late Modern British and Irish English. In *Contact, Variation, and Change in the History of English* [Studies in Language Companion Series 159], Simone Pfenninger, Olga Timofeeva, Anne-Christine Gardner, Alpo Honkapohja, Marianne Hundt & Daniel Schreier (eds), 239–256. Amsterdam: John Benjamins.

Sinclair, John. 1782. *Observations on the Scottish Dialect*. Edinburgh: W. Strahan, and T. Cadell, and W. Creech.

SND = Grant, William & Murison, David (eds). 1931. *The Scottish National Dictionary*. <http://dsl.ac.uk>

Taglicht, Josef. 1970. The genesis of the conventional rules for the use of *shall* and *will*. *English Studies: A Journal of English Letters and Philology* 51: 193–213. doi:10.1080/00138387008597374

Tieken-Boon van Ostade, Ingrid. 1985. 'I will be drowned and no man shall save me': The conventional rules for *shall* and *will* in eighteenth-century English grammars. *English Studies: A Journal of English Letters and Philology* 66: 123–142. doi:10.1080/00138388508598375

Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65(1): 31–55. doi:10.2307/414841

Visser, Fredericus Theodorus. 1969. *An Historical Syntax of the English Language*, Vol. 3. Leiden: E. J. Brill.

Warner, Anthony R. 1993. *English Auxiliaries: Structure and History*. Cambridge: CUP. doi:10.1017/CBO9780511752995

Appendix. Pearson's Chi-squared test results

Table 9. Chi-squared results for the changing distribution of modal meanings of *will* and *shall* between SC0/SC1 and SC2/SC3

	will	*shall*
volition	drop between SC0/SC1 and SC2/SC3 X^2 (1) = 44.8184, $p < 0.001$	no variation between SC0/SC1 and SC2/SC3 X^2 (1) = 0, $p = 1$
prediction	rise between SC0/SC1 and SC2/SC3 X^2 (1) = 32.5218, $p < 0.001$	random variation between SC0/SC1 and SC2/SC3 X^2(1) = 0.0187, $p > 0.05$
obligation	does not apply	drop between SC0/SC1 and SC2/SC3 X^2 (1) = 19.8988, $p < 0.001$
probability judgment	rise between SC0/SC1 and SC2/SC3 X^2 (1) = 13.1452, $p < 0.001$	random variation between SC0/SC1 and SC2/SC3 X^2 (1) = 0.3903, $p > 0.05$

Table 10. Chi-squared results for the changing distribution of modal meanings of first-person *will* and first-person *shall* between SC0/SC1 and SC2/SC3

	first-person *will*	first-person *shall*
volition	drop between SC0/SC1 and SC2/SC3 $X^2 (1) = 5.3364, p < 0.05$	random variation between SC0/SC1 and SC2/SC3 $X^2 (1) = 1.727, p > 0.05$
prediction	random variation between SC0/SC1 and SC2/SC3 $X^2 (1) = 3.1344, p > 0.05$	random variation between SC0/SC1 and SC2/SC3 $X^2 (1) = 1.2071, p > 0.05$

Table 11. Chi-squared results for the changing distribution of modal meanings of third-person *will* and third-person *shall* between SC0/SC1 and SC2/SC3

	third-person *will*	third-person *shall*
volition	drop between SC0/SC1 and SC2/SC3 $X^2 (1) = 47.9807, p < 0.001$	no variation between SC0/SC1 and SC2/SC3 $X^2 (1) = 0, p = 1$
prediction	rise between SC0/SC1 and SC2/SC3 $X^2 (1) = 24.0161, p < 0.001$	random variation between SC0/SC1 and SC2/SC3 $X^2 (1) = 2.1411, p > 0.05$
probability judgment	rise between SC0/SC1 and SC2/SC3 $X^2 (1) = 15.778, p < 0.001$	random variation between SC0/SC1 and SC2/SC3 $X^2 (1) = 0.1603, p > 0.05$
low-degree modality	random variation between SC0/SC1 and SC2/SC3 $X^2 (1) = 0.8693, p > 0.05$	rise between SC0/SC1 and SC2/SC3 $X^2 (1) = 41.5242, p < 0.01$

CHAPTER 10

A study of Old English *dugan*

Its potential for auxiliation

Kousuke Kaita
Chiba University, Japan

This study examines how the Old English preterite-present verb *dugan* 'to avail' developed into a modal auxiliary of ability ('can') in Middle English. Two factors in the Old English period play a role in this auxiliation process: (i) the influence of the modal precursor *magan* (meaning 'to be strong' as a lexical verb and 'can' as a modal) (> Modern English *may*) and (ii) the textual characteristics of medical texts. *Dugan* shares with *magan* a number of morphological and semantic properties. Furthermore, in medical texts, Old English *dugan* is attested as a lexical verb with the *to*-infinitive; *magan* as well is attested as a lexical verb and as a modal followed by the bare infinitive. It is argued that *dugan* starts combining with the bare infinitive because of the analogy with *magan* after the medical manuscripts were passed on to the Middle English period.

Keywords: preterite-present verb, modal auxiliary, auxiliation, infinitive, medical texts

1. Introduction

The development of modal auxiliaries from so-called preterite-present verbs (hereafter PPVs)[1] has attracted much scholarly attention (e.g. Lightfoot 1979; Birkmann 1987; Nagle & Sanders 1998). A noteworthy issue in this respect is the survival of six Old English (OE) PPVs[2] as Modern English (ModE) modals (e.g. *magan* > ModE

---

1. A list of abbreviations is given at the end of this study. The short titles, spellings, and line numbers of Old English texts are based on the *Dictionary of Old English Web Corpus* (*DOEWC*).

2. The six Old English PPVs that survived are: *āgan* (> ModE *ought*), *cunnan* (> *can/could*), *durran* (> *dare/dared*), *magan* (> *may/might*), *mōtan* (> *must*), and *sculan* (> *shall/should*).

may; *sculan*[3] > *shall*/*should*), against the loss of six other PPVs[4] (e.g. OE *unnan* was supplanted by *grant*, OE *þurfan* by *need*) (see Lightfoot 1979: 101–103; Plank 1984: 311–312). The OE PPVs that survived (such as *magan* and *sculan*) were categorized "pre-modals", while the ones that were lost (such as *dugan*) were labeled "non-pre-modals" (see Lightfoot 1979: 101). According to Plank (1984: 312), of the whole group of PPVs, the pre-modals survived in the history of English because, as a well-structured system of grammatical elements, they "do not normally admit of extensive synonymy". Both *dugan* and *magan*, for instance, "had dynamic-modal meaning, referring to physical capability", but *dugan*, which conformed to common inflectional patterns, was lost before the ModE period. According to Nagle and Sanders (1998: 253), however, *dugan* "survives as a modal in some regional varieties of British English".

Thus far, there have been few studies on how *dugan* became a modal between OE and Middle English (ME). This paper, therefore, aims to discuss what allowed OE *dugan* to become a modal auxiliary between OE and ME. The hypothesis of this study is that the auxiliation of *dugan* occurred in the ME period, but that this process reflects the influence in the OE period (i) of the morphosyntactic and semantic analogy with *dugan*'s modal precursor *magan* and (ii) of textual characteristics. With regard to the latter, *dugan* shows several morphosyntactic usages – similar to those of *magan* (and secondarily *sculan*) – that are specific of one genre in particular, namely medical texts. The medical texts (e.g. *Leechbooks*; see Krischke 2013), have their own tradition that they handed over to the early ME period. These lexical and textual properties may have led to the auxiliation of *dugan* from ME on.

To demonstrate the development of *dugan* as influenced by *magan* (and partly *sculan*), I shall first point to the fact that the PPVs *dugan* and *magan* share the same lexical origin (Section 2). After reviewing how a lexical verb changes into a modal (Section 3), the usages of OE *dugan*, *magan*, and *sculan* will be analyzed in Section 4. Section 5 briefly mentions the subsequent history of OE *dugan* (ME *douen*) in the ME period and in the Scots dialect, which attests *dow*. Additional support for my argumentation will come from cognate verbs of OE *dugan* in older Germanic languages (see Section 6). Finally, Section 7 summarizes the influence of OE *magan* on *dugan* and concludes the argument.

---

3.   The infinitive form of *sculan* in OE is not attested, so it would be more accurate to asterisk it as **sculan*. The *DOEWC* records the infinitive *dugan* as an emended form in *RevMon* (Whitelock) 274 (see also Hogg & Fulk 2011: 301).

4.   They are: *dugan* 'to avail', *genugan* 'to suffice', *þurfan* 'to need', *munan* 'to remember', *unnan* 'to grant', and *witan* 'to know'.

2.   Old English *dugan*: Its etymology

Etymologically, OE *dugan* and *magan* share the same unusual morphological origin in Proto-Indo-European (PIE). To demonstrate this, I will first describe the typical formation of PPVs, as exemplified by *sculan*. I will then discuss the atypical formation of *dugan* and *magan*, in comparison to that of *sculan*.

Germanic PPVs are commonly viewed as verbs whose (reduplicated) perfect conjugational forms in PIE are reanalyzed as present forms, as a result of which new dental weak preterite forms were created in Proto-Germanic (PGmc) (see Birkmann 1987 for detailed discussion). The development of Gothic *skulan* 'to owe, must' (cognate to OE *sculan*) is a case in point. Its PGmc forms were the present *skal* (Gothic *skal*; OE *sceal*) and the preterite *skuldē* (Gothic *skulde*; OE *sceolde*) (see Ringe 2006: 261). The former stems from PIE *(s)ke-(s)kól/(s)k̥l-* (see *LIV2*, s.v. ?*(s)kel-* "schuldig werden" ('to become guilty')), the reduplicated perfect form of PIE *(s)kél-e-*; the latter is the PGmc dental preterite form.

Unlike *sculan*, the PGmc ancestors of OE *dugan* and *magan* are not based on the PIE perfect forms, but on the stative present forms. Gothic *daug* (cognate to OE *deah*) stems from the stative present form *$d^h u g^h$-éi̯* in PIE (< root *$d^h e u g^h$-* "treffen" (*LIV2*) ('to be fitting, suitable')). Similarly, Gothic *mag* (cognate to OE *mæg*) can be traced back to the stative present *$mág^h ei̯$* in PIE (< root *mag^h-* "können, imstande sein" (*LIV2*) ('can, to be able to')).[5] Consequently, OE *dugan* and *magan* share a similar and exceptional formation in Germanic.

3.   Auxiliation of pre-modals

Dugan and *magan* show a similar lexical origin (compared to *sculan*), and *dugan* also follows *magan* formally and semantically in its auxiliation process. Recall that auxiliation, as Kuteva (2001: 1–2) defines it, refers to "a morphosyntactic change whereby the lexical structure (1) *verb – complement* turns into the grammatical structure (2) *grammatical marker – main verb*". Modal auxiliary constructions, in particular, contain verbs that originally were lexical verbs at stage (1) and changed into grammatical markers in stage (2) expressing deontic, dynamic, or epistemic modality. This section briefly outlines the different formal and semantic changes a lexical verb (in particular, a pre-modal) undergoes in becoming a modal auxiliary. Special attention will be given to the auxiliation of *magan* and *dugan* vs. *sculan*.

---

5.   The spellings of PIE stative present forms are from Kümmel (1996: 62, 80), and those of the root forms are from *LIV2* (s.v. *$d^h e u g^h$-*, *mag^h-*).

3.1 Some formal features of the modals

According to Quirk et al. (1985: 137), ModE central modals (e.g. *shall*, *should*, *may*, *might*, etc.) share the following criteria: (i) they are used with the "[b]are infinitive" (Quirk et al.'s example: *I can go*); (ii) they have "[n]o nonfinite forms" (**to can/*canning/*canned*); (iii) they have "[n]o -*s* form" (**She cans come*); and (iv) they have "[a]bnormal time reference" (*You could leave this evening* "[not past time]"). The question, then, is: what changes did the OE PPVs, and in particular *sculan*, *magan*, and *dugan*, undergo in the auxiliation process?

   The OE PPVs *sculan* and *magan* and their ME correlates *shulen and mouen* (SHALL and MAY in Warner 1993) acquired the formal features of auxiliation, as distinguished by Quirk et al. (1985), at a different pace. While *sculan/shulen* as well as *magan/mouen* are used with bare infinitives (pertaining to feature (i) in Quirk et al.) and do not show the third-person singular -*eþ* suffix that is the precursor of -*s* (feature (iii)) (see Lightfoot 1979: 101–102; Fischer 2003: 18),[6] it is only *sculan/ shulen* that has finite forms only (feature (ii); see Warner's 1993: 153 "[p]roperties of verbs in Old and Middle English") and that uses "past-tense forms without past time reference, outside a motivating context" (Warner 1993: 152; feature (iv)). For *magan/mouen*, Warner's (1993: 145) table "[r]ecorded nonfinites of the major preterite-present verbs and* will" (feature (ii)) demonstrates that there were infinitival and present-participle forms in the OE period, and that past-participle forms were also added in the Late ME period; furthermore, past-tense forms without past-time reference only developed from ME onwards (Warner 1993: 153; feature (iv)).

   The fact that the auxiliation pace was not simultaneous for OE *sculan* and *magan* raises the question whether *dugan* followed the development of *sculan* or of *magan*. *Dugan/douen* is found to be similar to *magan/mouen*. With regard to feature (i) in Quirk et al. (1985), ME *douen* comes to be used with bare infinitives possibly imitating the structure of *mouen*. As for feature (ii), Warner's (1993: 145) table of nonfinite forms shows that OE *dugan* and *magan* share properties in that infinitival and present-participle forms are attested for both, while past-participle forms are not; in ME, however, all three of these nonfinite forms are attested for *douen* and *mouen*. As for feature (iii), both *dugan/douen* and *magan/mouen* lack -*s* endings. With regard to (iv), in their ability sense ('can, be able to'), the past-tense forms and the time reference of *douen* and *mouen* are not always consistent (see examples (15)–(17)).

---

6. In contrast to the absence of the -*eþ* ending, *sculan/shulen* and *magan/mouen* have -*t* in the second person singular (e.g. *scealt* in Examples (1) and (2) and *meaht* in (4)).

3.2   Evaluative readings

In their auxiliation process, PPVs also undergo semantic changes from a lexical to a modal meaning. Here as well, *dugan* is seen to follow a path similar to that of *magan*, in that it is synonymous with *magan* as a lexical verb and that *dugan* and *magan* both later become dynamic modals of ability 'can'; *sculan*, on the other hand, became a deontic modal expressing obligation 'must'.[7]

An important concept in discussing the semantic change of *sculan*, *magan*, and *dugan* is that of speaker-related (or subjective) evaluation. These verbs may involve a speaker-related, evaluative (appreciative or pejorative) component as part of their semantic content (as do items such as *snug*, *idiot*), or they may encode nothing but speaker evaluation with respect to a state of affairs (as does *may* expressing epistemic stance with respect to a state of affairs) (see Cuyckens et al. 2010: 8).

Let us first consider the auxiliation of *sculan* in OE. *Sculan* is used as a lexical verb in its meaning of 'to owe', as in (1), or as a deontic modal of obligation 'must', as in (2).

(1)   *Agyf þæt   þu   me   **scealt**.*
     pay   what   you   me   owe
     'Pay what you owe me.'   (*Mt* (*WSCp*) 18.28, s. xi[1]–xii)[8]

(2)   *Ageld nu   swiðe raðæ   þæt þæt ðu   me geldæn   **scealt**.*
     pay   now   very   quickly   that   that   you   me   pay   must
     'Pay now very quickly what you must pay me.'   (*ÆHomM* 7 (Irv 2) 20, s. xii[2])

In (1), *sculan* describes a situation in which someone owes money to someone else. In OE example (2), of a later date, *sculan* accompanies the main verb *geldan* 'to pay'. The form *scealt* encodes the speaker-related evaluative meaning of obligation 'must', which derives from the fact that a person A who is in debt to another person B is in a state of obligation to the other person B, i.e. A must pay money back to B.

Magan, another OE PPV, is also attested as a lexical verb, as in (3).

(3)   *hit ne   **mæg**   syððan to   nahte*
     it   not   is.good   then   for   nothing
     'it (= salt of the earth) is then good for nothing'   (*Mt* (*WSCp*) 5.13, s. xi[1]–xii)

---

7.   Based on Palmer (2003: 7), I associate the meaning of ability 'can' with dynamic modality, and that of obligation 'must' with deontic modality.

8.   In the examples cited, the emphasis, glossing, and translations are mine unless otherwise noted. The dating of the OE examples cited is based on the manuscript dating described by Ker (1957).

This example is recorded in *BT* (s.v. *magan*, I. "*to be strong, efficacious, to avail, prevail, be sufficient*", Magan tó: "*to serve a purpose, be good for, have an effect, be the cause of*").[9] *Magan*'s semantic content as a lexical verb comprises such positive notions as 'strong', 'efficacious', or 'good', and thus involves the speaker's evaluative (subjective) judgment. After *magan* has become a dynamic modal of ability ('can') as in (4), it refers to the positive potential that its subject (*þu*) carries out the proposition (*singan*).

(4) *Hwæðre   þu   **meaht** singan.*
     nevertheless you can   sing
     'Nevertheless you can sing.'   (*Bede* 4 25.342.31, s. x[1])

OE *dugan* appears as a main verb 'to be useful', glossing Latin *prodesse* 'to profit' in (5). No example is attested of its use as a modal (with the bare infinitive).

(5) [*quid enim   **prodest** homini* […]]
     *huæt forðon   **deg**   menn* […]
     what therefore profits   man
     'Then, what profits a man […]?'   (*MtGl* (*Li*) 16.26, s. x[2])

DOE (s.v. *deag*, A.2.) defines *deg*, as in (5), as "avails, is helpful; often with dative object: is helpful for/to (someone), benefits (someone), does (someone) good". In other words, *deg* in (5) means that the subject *huæt* is regarded as profitable to a person in the dative case (*menn*). In this sense, the meaning of *dugan* shares its positive evaluative (speaker-related) meaning with that of *magan* 'to be strong', rather than with the meaning of *sculan* 'to owe', which is not speaker-related.

To summarize the discussions on *sculan*, *magan*, and *dugan* at the level of the lexical verb, OE *dugan* and *magan* are semantically similar in that both convey a positive evaluative, speaker-related judgement, thus differing from *sculan*, whose lexical meaning is not subjective (in Traugott's 1989: 34 terms, *sculan* expresses a meaning "based in the external described situation"). The next section will discuss the usage of OE *dugan* in detail and compare with *magan* and *sculan*, wherever relevant.

---

9. Old Saxon *daug* (cognate to OE *dugan*) is used in this very same context in *Heliand* (see Section 6).

4. Old English *dugan*

In this section, I will show that *dugan* in OE is used as a lexical verb in various text genres, conveying a speaker-related positive evaluation. In particular, *dugan* is seen to occur prominently in medical prose texts. Its usage will also be argued to follow that of *magan* and *sculan*. *Dugan* does not function as a modal in OE because it takes no bare infinitive, but this OE stage provides *dugan* with the auxiliation potential, and *dugan* starts auxiliation in the ME period.

I collected 114 tokens of *dugan* from the *DOEWC*. The forms retrieved are: *dæg*; *deag*; *deah*; *deahg*; *deg*; *dege*; *dohte*; *dohten*; *dohtest*; *dohton*; *dugan*; *duge*; *dugende*; *dugon*; *dugunde*; *duhte*; *dyge*.[10] These tokens are found in various text styles and genres, showing the distribution presented in Table 1.

Table 1. Distribution of *dugan* in the *DOEWC*

Style	Genre	Number of occurrences
1. Verse	Heroic poems (*Beo*; *Mald*; *Wid*)	12
	Religious poems (*And*; *GuthA*; *ChristA*; *GenA*; *GenB*; *Sat*; *Prec*)	9
	Riddles (*Rid* 61/73)	2
	Others (*Met*; *MCharm* 11)	2
2. Prose	Medical texts (*Lch* I (Herb); *Lch* II (1)(2); *Med* 3/5.2/5.10; *PeriD*)	51
	Prognostics (Dreambooks) (*Prog* 3.2/3.10/6.8)	9
	Homilies (*ÆAbus* (Mor)/(Warn); *HomU* 25/40; *WHom* 20.1/2/3)	9
	Chronicles (*ChronC/D/E*)	6
	Others (*Bo*; *Deut*; *ÆLS* (Alban); *ÆLet* 1 (Wulfsige Xa); *Let* 2 (Clayton); *Ch* 1188 (HarmD 1); *RevMon* (Whitelock))	11
3. Gloss	*MtGl* (Li); *DurProv*	3
Total		**114**

---

10. Listed in alphabetical order, whereby grammatical function (e.g. indicative/subjunctive, singular/plural, etc.) is disregarded. I excluded from these 114 examples the case of *gedige* in *MCharm* 9 13: *Gif hyt hwa gedo, ne gedige hit him næfre* 'If someone does it, it will never be of benefit to him'. *Gedige* in this sentence (*DOE*, s.v. *gedēag* "avails") is synonymous with *dugan*. *BTS* (s.v. *ge-dígan*) has two definitions: (i) "*to escape* danger"; (ii) "*to benefit, profit*", which contains the example above. *Gedigan* is regarded as a derivative verb from *dugan*. *LIV2* (s.v. *$^*d^heugh$-*) defines OE *gedȳgan* as "bestehen, überwinden" ('to endure, overcome'), and regards it as derived from PIE *$^*d^hough$-éie-*, a causative form of *$^*d^heugh$-*. Although synonymous with *dugan*, I omit this secondary verb *gedigan* from the present investigation.

As can be seen from the distribution in Table 1, the majority of Examples (51 out of 114) appear in medical prose texts. The following three Subsections (4.1–4.3) present characteristic examples of *dugan* according to textual genre.

4.1   Verse texts

Several examples of the lexical verb *dugan* in verse texts occur in formulaic expressions. In each of these texts, the author assigns a positive evaluation to the subject of *dugan*, in that its courage is useful or effective. One such example is (6), taken from *Beowulf*; it consists of a conditional conjunction *þonne*, the subject *his ellen*, and the verb *dugan* denoting that "his courage" is effective.

(6)   *þonne his ellen   **deah***
       when   his courage is.effective
       'when his courage is effective'   (*Beo* 572, s. x/xi)

Formulae like (6) are also found in *Beo* 589, *And* 460, *Sat* 282, and *Rid* 73 9,[11] and are discussed in many studies.[12] The OE poem *Andreas*, for instance, contains a formula *gif his ellen deah* 'if his courage is useful'. *Andreas* is thematically biblical, but the religious theme is adapted to a Germanic narrative. Andrew is here depicted as a valiant warrior on a heroic expedition. The description of his courage plays a crucial role, and this type of formula functions as an effective device in illustrating Andrew's valor, reminiscent of such heroic poems as *Beowulf*.

4.2   Homiletic prose

Dugan also predicates the author-related notion of effectiveness of its subject in homiletic prose. It has a particular wording in homilies that is different from that in medical texts. An example is (7), taken from a homily by Wulfstan.

(7)   *Ne **dohte**   hit nu   lange inne   ne   ute*
       not prospered it   now long within nor without
                                              (*WHom* 20.3 55, s. xi (3rd quarter))
       'Nothing has prospered now for a long time, within or without.'
                                                           (Bethurum 1957: 360)

---

11. In addition, there is a similar expression in Old High German *Hildebrandslied* (see Section 6).

12. See Ker (1908: 87), Chadwick (1912: 78), Dickins (1915: 83), Clubb (1925: 87), Dobbie (1953: 146), Brooks (1961: 77), Cavill (1999: 144–146), and Fulk et al. (2008: 153).

As Bethurum (1957: 360) indicates, a particular author evaluation is predicated of the subject (*hit*). Here, the evaluation is negative because of the negative particle *ne*. This construction – with almost the same wording in the negative – appears seven times in Wulfstan's various homilies,[13] and, as pointed out by Bethurum (1957), a similar passage occurs in the *Peterborough Chronicle* (*ChronE*) (see example (8)).

(8) *þet   him   naðor   ne   **dohte**   ne   innhere   ne   uthere*
so.that to.them neither not did.good nor native.army nor foreign.army
'so that neither the native army nor the foreign army did good to them [= the land people]'   (*ChronE* (Irvine) 1006.10, s. xii[1], xii med.)

Given the widespread use of these formulaic expressions involving *dugan* in homiletic texts (and also verse texts, as mentioned in Section 4.1), it appears that the evaluative interpretation of *dugan* is conventionalized in OE.

4.3   Medical prose

The majority of occurrences of the PPV *dugan* are attested in several manuscripts containing medical texts, which are listed in Table 2.

Table 2. Medical texts examined

Short title in *DOEWC*	Manuscripts	Dating*
Med 5.10	Louvain-la-Neuve, Centre Général de Documentation, Université Catholique de Louvain, Fragmenta H. Omont 3 (Ker Supplement 417; Gneuss-Lapidge 848)	s. x in.
Lch II (1)(2)(3)	London, British Library, Royal 12. D. xvii (Ker 264; Gneuss-Lapidge 479)	s. x med.
Med 5.2	London, Wellcome Library for the History and Understanding of Medicine, 46 (Ker 98; Gneuss-Lapidge 523)	s. x/xi
Med 3; *Lch* I (Herb); *Med* 1.1	London, British Library, Harley 585 (Ker 231; Gneuss-Lapidge 421)	s. x/xi, xi[1]
Lch II (Fragment)	London, British Library, Harley 55, fols. 1–4 (Ker 225; Gneuss-Lapidge 412)	s. xi[1]
Lch I (Herb); *Med* 1.1	London, British Library, Cotton Vitellius C. iii, fols. 11–85 (Ker 219; Gneuss-Lapidge 402)	s. xi[1]
	Oxford, Bodleian Library, Hatton 76 (S.C. 4125), fols. 68–139 (Ker 328; Gneuss-Lapidge 633)	s. xi med.
PeriD	London, British Library, Harley 6258B	after 1200 (Ker: p. xix)

* Based on Ker (1957).

---

13. Similar occurrences to (7) are: *HomU* 25 (Nap 27) 14; *HomU* 40 (Nap 50) 78; *WHom* 20.1 50; 20.2 63; 20.2 106; 20.3 109.

The manuscripts in Table 2 are dated within the transitional period between OE and ME.[14] According to Hollis (2001: 198) for ᴍs Royal 12. D. xvii, "[t]he exemplar is dated *c*.900 on linguistic grounds (Quirk 1955)". Note also Hollis' (2001: 204) reference to the use and transmission of some of the manuscripts mentioned in Table 2:

> Royal 12 D. xvii and Harley 585 continued to be used and added to throughout the eleventh and twelfth centuries, and so also were the translations of the *Herbarium* in Vitellius C. iii and Hatton 76.

As OE medical texts have social and cultural significance, and contain many tokens of *dugan* (51 in total, 40 of which in *Lch* (1)(2) alone), there is considerable reason to assume that the usage of *dugan* as characteristic of the medical texts has affected its later linguistic development.

As in the text types previously discussed, the OE PPV *dugan* in medical texts is used as a main/lexical verb. In 23 cases, it is accompanied by the preposition *wið* 'against', as in (9). The object of *wið* is usually the name of a disease or unrest, and the typical case of "*dugan* + *wið*" in *Lch* II describes an efficacious remedy (*dugan*'s inanimate subject) against some trouble or disease.

(9) *hit eac* **deah**   *wiþ*   *feondes costungum yflum*.
  it also is.efficacious against fiend's temptings evil
                                 (*Lch* II (2) 65.3.11, s. x med.)
  'it [= myrrh rubbed into wine] also is efficacious against the evil temptings of the fiend.'   (Cockayne 1865: 295)

As well, *dugan* in main verb use takes the *to*-infinitive in adjunct function in the medical texts *Lch* II (2) 30.1.6, *Lch* II (2) 34.2.1 (Example (10)), and *Med* 3 (Grattan-Singer) 10.1. The denotation of *dugan* in (10) is again 'good' or 'effective', and the connotation is positive. There is no example of *dugan* with a bare infinitive.

(10) *Wiþ*   *latre meltunge, olisatrum hatte wyrt*, *seo* **deah** *to*
  against late digestion olustratum called wort which is.good to
  drincanne.   (*Lch* II (2) 34.2.1, s. x med.)
  drink
  'For late digestion; a wort hight olustratum, which is good to drink.'
                                 (Cockayne 1865: 239)

OE *magan* as well has a main verb usage which is found frequently (21 times) in medical texts. Just like *dugan* in (9), *magan* in (11) takes an inanimate subject and

---

14. The use of medical manuscripts has its own tradition in the textual history of medieval English; see Mäkinen (2004: 149–152, 171), who describes the textual tradition from Greek and Latin in detail. I am not sure if *dugan* is employed to translate certain concepts from Greek or Latin sources in the medical texts in Table 2.

means 'to have power' or 'is valid', followed by the preposition *wið* ("*to prevail with or against, to be efficacious against* (of a medicine) *to be good for* (*a disease*)" in *BT*, s.v. *magan*; "[m]ay *as an independent verb*" in Visser 1963: 162–163).

(11) *Læcedom wiþ þære geolwan adle   &   wið þæm miclan lice   &*
  leechdom for   the   yellow   disease and for the   mickle body and
  *dolhdrencas   twegen &   oþer   **mæg**   wiþ lungen wunde eac.*
  wound.drinks two   and another is.valid for lung   wound also

  (*Lch* II (2 Head) 61, s. x med.)

  'A leechdom for the yellow disease *or jaundice*, and for the mickle body *or elephantiasis*; and two drinks for wounds, one of which is valid for a lung wound also.'   (Cockayne 1865: 173: additional explanations by the translator)

In the medical texts, the inanimate subject may have undergone some kind of personification because the contextual importance lies in the subject (a drink in (11)) being strong and effective, as if the thing were a person.

In addition to *dugan* and *magan*, *sculan* collocates with the preposition *wið*, as shown in (12).

(12) *Wið genumenum mete genim elehtran, lege under weofod, sing nigon*
  for taken   meat take   lupins   lay under altar   sing nine
  *mæssan ofer, þæt **sceal** wiþ genumenum mete*
  masses over that shall for taken   meat

  (*Lch* II (1) 67.1.1, s. x med.)

  'For *the better digestion of* meat taken; take lupins, lay them under the altar, sing over them nine masses, that shall avail for meat taken'

  (Cockayne 1865: 143: additional explanations by the translator)

The subject of *sculan* in (12) is *þæt* 'that' referring to the whole transaction. This usage appears to be different from the lexical verb usage of *sculan* meaning 'to owe' or 'to have money to repay', as exemplified in (1). The context in (12) evokes no hint of financial debt. Visser (1963: 162) terms this use of *sculan* as "[i]ndependent shall" meaning "'is proper', 'is due', 'is to be given, applied, done'". I would like to argue that this independent use is likely the result of a syntactic analogy with *magan* with the preposition *wið*.

In *Lch* II, *magan* and *sculan* are also attested as modals of ability and obligation, respectively, as exemplified in (13). The subject of *magan* and *sculan* in (13) is the indefinite pronoun *man*. This indicates with regard to the descriptive and instructive style of the medical texts that there is little (or no) need to specify the agent who carries out the proposition. That is, in (13), the focus is on understanding and treating the sore, rather than on the (unspecified) person who deals with the sore.

(13)   *Her   sindon   læcedomas   wiþ   æghwæþerre   sidan   sare   &   tacn*
　　　 here   are   　leechdoms   against   of.either   　side   sore   and   tokens
　　　 *hu   sio   adl   　 toweard   sie   &   hu   þæt   mon   ongitan   **mæge***
　　　 how   the   disease   toward   is   　and   how   that   one   understand   can
　　　 *&   hu   　hiora   mon   tilian   **scyle**. 　　　 (*Lch* II (2) 46.1.1, s. x med.)
　　　 and   how   it   　 one   treat   must
　　　 'Here are leechdoms for sore of either side, and tokens how the disease ap-
　　　 proaches, and how a man may understand that, and how a man shall treat it.'
　　　　　　　　　　　　　　　　　　　　　　　　　　　　　(Cockayne 1865: 257)

The interplay between *magan* and *sculan* can be summarized as follows. Both are
used as modals of ability and obligation. Although *magan* and *sculan* are seman-
tically different, they share syntactic properties: collocation with a bare infinitive
and occasional collocation with a *man*-subject in the modal uses. *Magan* also has
a main verb usage with the preposition *wið*. This usage of *magan* (as a main verb
and modal) may have given rise to *sculan*'s independent usage.

　　　Table 3 shows how *dugan* relates to *magan* and *sculan* in the two parts of the
Lch II manuscript.[15] As can be seen, *dugan*, *magan*, and *sculan* behave differently
in these two parts. While *Lch* II (1)(2) comprises many attestations of main verb
dugan, *Lch* II (3) presents no occurrences. Furthermore, *magan* and *sculan* display
main verb/independent usages as well as modal verb usages in *Lch* II (1)(2); in
contrast, in *Lch* II (3), *magan* is used only once as a main verb, while *sculan* is used
exclusively as a modal.

Table 3. OE *dugan*, *magan*, and *sculan* in *Lch* II (1)(2)(3) (normalization per 10,000 words)

		Lch II (1)(2)		*Lch* II (3)		Total
		Absolute figures	**Normalized figures**	**Absolute figures**	**Normalized figures**	
dugan	(Main verb)	40 12.82%	11.9	– –	–	40
magan	(Main verb)	6 1.92%	1.8	1 1.92%	1.8	7
	(Modal)	128 41.03%	38.2	38 73.08%	68.8	166
sculan	(independent use)	6 1.92%	1.8	– –	–	6
	(Modal)	132 42.31%	39.4	13 25.00%	23.6	145
Total		**312 100.0%**	**93.1**	**52 100.0%**	**94.2**	**364**

---

15. *Lch* II (1)(2) contains, according to the word count in *DOEWC*, 33517 words in total, and
Lch II (3) contains 5520 words.

Table 4 presents the syntactic distribution of *dugan*, *magan*, and *sculan* in the following medical texts: *Med* 5.10; *Lch* II (1)(2)(3); *Med* 5.2; *Med* 3; *Lch* II (Fragment); *Lch* I (Herb); *Med* 1.1 (Cotton Vitellius C. iii) (77872 words in total).[16]

Table 4. The syntactic collocation of *dugan*/*magan*/*sculan* in OE medical texts

Usage	collocation	*dugan*		*magan*		*sculan*	
Main verb (*dugan* & *magan*) or independent use (*sculan*)	Ø	6	12.2%	1	0.3%	2	0.9%
	dative	18	36.7%				
	to (preposition)	2	4.1%			1	0.5%
	wið	23	46.9%	20	7.0%	6	2.7%
Modal (*magan* & *sculan*)	*man*-subject			77	26.9%	116	52.3%
	other subjects			188	65.7%	97	43.7%
Total		**49**	**100.0%**	**286**	**100.0%**	**222**	**100.0%**

An important share of the *dugan* examples appear in collocation with *wið*. *Magan* is used as a lexical verb with *wið* and also as a modal. Although a few attestations of *sculan* are found with *wið*, it is used almost exclusively as a modal. That is, among these three PPVs, *dugan* and *magan* are similar in describing the effectiveness of the subject against certain trouble indicated by *wið*. This evaluative interpretation in medical texts, together with its usage with the *to*-infinitive in adjunct function (as in example (10)), provided *dugan* with the potential to be auxiliarized in the subsequent period.

The OE period thus yielded the conditions facilitating the auxiliation of *dugan*. It appears that a positive evaluative interpretation of OE *dugan* (typically in formulaic expressions) was found in such specific genres as verse (e.g. *Beo*) and homilies (e.g. *WHom*). In medical texts, *dugan* often occurs in the structure of "PPV + *wið*" (beside *magan* and *sculan*), meaning 'is good, is effective'. *Dugan* in this text genre is also present as a lexical verb in combination with the *to*-infinitive (with adjunct function). *Magan* and *sculan*, as modal verbs, take the bare infinitive. Despite different infinitival types, *dugan* thus shows a syntactic and semantic similarity to *magan* (and *sculan*), and it can thus continue to tread along the pathway to auxiliation as the medical manuscripts in OE were transferred to the ME period. The medical texts, therefore, played a significant role in the linguistic change of *dugan*.

---

16. Recall that in Table 1, 51 occurrences of *dugan* were mentioned. The total count of *dugan* in Table 4 is 49, as it excludes one occurrence of *dugan* in the present-participle form in *Lch* II (1) and one occurrence of *dugan* in *PeriD* which belongs to the ME period (see Ker's 1957: xix dating in Table 2).

5. Middle English and Modern English periods

This section describes the subsequent history of the auxiliation of *dugan* (ME *douen*), based on an analysis of the examples collected from the *OED* and Visser (1969: 1444–1445). I will argue the following points: (i) *Douen* is used with the *to*-infinitive in Early ME, and then with the bare infinitive from the Early Scots period on, as is ME *mouen* (< OE *magan*); (ii) Following *mouen*, the notion of effectiveness is converted into the notion of ability; (iii) The impersonal usage of *douen* is reanalyzed as personal usage; (iv) The syntactic and semantic similarity of *douen* and *mouen* remains valid in the Northern dialect in ME. In sum, the auxiliation of *douen* is achieved within the ME period as it is continuously affected by *mouen*.

5.1 Middle English *douen*

Consider Example (14) from *St. Juliana* (D'Ardenne 1961: 45), recorded in ᴍꜱ Bodley 34;[17] it instantiates the so-called impersonal verb construction in the Early ME period.

(14) *as meiden* **deh**   *to beonne*
       as maiden is.proper to be
       'as it is proper for a maiden to be'
                                   (*St. Juliana* 487 (ᴍꜱ Bodley 34), c1225(?c1200))[18]

Evidence for the impersonal status of construction in (14) comes from (i) the lexical entry of *deh* in the *OED* ("s.v. *dow*, †4. [t]o be good, fitting, or proper for any one; to become, befit, behove. Usually impers[onal]" (examples: *a*1225–*c*1450)); (ii) the dative case form of the noun *meiden*. Visser (1969: 1444) cites the example (14) as the first case of ME *douen* (given as *dowen*) taking an infinitive and argues that "[i]n the beginning *dowen* was often used 'impersonally'", and *douen* used to have a pronominal (*h*)*it*-subject and (*h*)*it* was then lost. Following this, "the noun originally acting as an indirect object has lost the dative endings" (p. 1444), and "the 'impersonality' of the verb is obscured, and these nouns tend to be realized as the subjects of the utterances" (p. 1444).
     Semantically, the pathway of *dugan*/*douen* within OE and ME can be described as follows, with (i) referring to the interpretation of OE *dugan*, and (ii) and (iii) to that of ME *douen*:

---

17. Oxford, Bodleian Library, ᴍꜱ Bodley 34.

18. The dating of the ME examples cited is based on that provided in *MED*.

(i) usefulness: '[subject] is useful (for *x*) (+ *to*-infinitive)'

(ii) propriety: 'it is proper (for *x*) (+ *to*-infinitive)'

(iii) a. ability: '*x* is able *to*-infinitive'
b. obligation: '*x* is obliged *to*-infinitive'

With regard to (i), OE *dugan* as a main verb denotes usefulness applying to its nominative subject, sometimes accompanied by a dative noun *x*. When *dugan* also collocates with a *to*-infinitive (as in example (10)), the infinitival proposition can be interpreted as the activity the subject is useful for. The notion of "usefulness" associated with *dugan* expresses a positive connotation, thus giving rise to the reading that it is 'proper' to carry out the infinitival proposition, as described in (ii); this usage is instantiated by ME *douen* in (14).

Along with the syntactic reanalysis of the noun *x* in the dative case to the nominative (see Visser 1969: 1444), the further pathway for *dugan* with an infinitive is either (iiia) the dynamic modality of ability (as it is also conveyed by *mouen* 'can'), or to (iiib) the deontic modality of obligation (as it is also conveyed by *shulen* 'must'). The deontic reading of (iiib) is confirmed by another manuscript reading of (14). In MS Royal 17. A. xxvii,[19] *ouen* (*ah*) meaning 'ought' (< OE *āgan* 'to have' with inflected infinitive) is used for *douen* (*deh*) and the context would be translated as 'as a maiden ought to be'.[20] The noun *meiden* in this case can be interpreted as a nominative subject.[21] The shift from ME *douen* to an 'ability' reading, then, proceeds by analogy with *mouen* because their semantic affinity had been prevalent since the OE period (or even perhaps since the PIE period as indicated by their characteristic formation).

---

19. London, British Library, MS Royal 17. A. xxvii. The text (D'Ardenne 1961: 44) reads: *as meiden ah te beonne*.

20. The pathway from (ii) to (iiib) may also be confirmed by looking at some uses of the verb *behoven* 'to behove' in ME *Cursor Mundi* (MS Cotton, see footnote 22): line 1665 *A schippe be-houes þe to dight* 'It is proper for you to build a ship'; lines 9797–9798 *Qua suld ma þis ranscun þan / Behoued bath be godd and man* 'Whoever should make this ransom then, it behoved to be both God and man'. *Be-houes* in 1665 is defined in the *MED* (s.v. *bihoven* 2b. (b) "with respect to custom, morals, doctrine, office, etc.: it is incumbent upon him (to do sth., etc.); […] it is proper or fitting for him"). For line 9798, another MS Oxford, Bodleian Library, Fairfax 14 attests a modal *must*: *hyt must be bothe god and man* 'It must be both God and man', which is a context of obligation.

21. While D'Ardenne's (1961: 85) glossary assigns the label "impers[onal]" to *deh*, there is no such remark for *ah*.

The lexical affinity between *douen* and *mouen* continues further. In Example (15), from ME *Cursor Mundi* (ᴍꜱ Cotton)[22] (Morris 1874–1892: 1358), both *douen* and *mouen* are located in the end rhyme, and both mean 'was able to' or 'could'.

(15) *Fight he aght   ai   quils   he **dught**,*
      fight he ought always while he was.able.to
      *And fle   quen   he langer ne   **moght**.*
      and flee when he longer not could
      'He ought to have always fought while he was able to, and fled when he could no longer (fight).'   (*Cursor Mundi* 23771–23772 (ᴍꜱ Cotton), a1400(a1325))

Example (15) *dught* here is the first example in the *OED* s.v. *dow*, 5.: "[t]o have the strength or ability, to be able (to do something)" (examples: *a*1300–1818)). The analysis of *douen* in (15) as an auxiliary can be confirmed with respect to the rhyming affinity between *douen* and *mouen*.

The following pair (16) and (17) in *Cursor Mundi* (1874–1892: 618–619) exhibits the lexical alternation between *douen* and *mouen* in two manuscripts of the Northern dialect: ᴍꜱ Cotton and ᴍꜱ Göttingen.[23]

(16) *Quen ioseph sagh na hide   ne   **dught***
      when Joseph saw   no hiding not availed
      'When Joseph saw (that) no hiding availed'
                                    (*Cursor Mundi* 10771 (ᴍꜱ Cotton), a1400(a1325))

(17) *Quen ioseph þat   he ne   **miht**   noght*
      when Joseph that he not could not
      'When Joseph (saw) that he could not (hide)'
                                    (*Cursor Mundi* 10771 (ᴍꜱ Göttingen), a1400)

Morris (1892: xlvi), translating (16) as "[w]hen Joseph saw that no secrecy (hiding) availed", comments on the Göttingen version: "G's text leaves out *na hide*, and turns *ne dught* into *ne might noght*". These three examples of (15)–(17) from *Cursor Mundi* demonstrate the close lexical (and semantic) relation between *douen* and *mouen*, although the syntactic difference between *douen* in (16) (a main verb) and *mouen* in (17) (a modal) should not be overlooked. In the next section, the subsequent period in the northern dialect will be discussed.

---

22. London, British Library, ᴍꜱ Cotton Vespasian A. iii.

23. Göttingen, University Library, ᴍꜱ Theol. 107.

5.2   Scots *dow* 'can'

There are some examples of Scots *dow* with the modal meaning 'can'. The first example of *dow* (see (18)) in Early Scots with the bare infinitive recorded in *DOST* can be traced back to the 14th century (s.v. *dow*, 2.a. "[t]o be able or fit to do something"; "[w]ith simple infinitive; = 'can', 'could'"). The infinitive *chang* is used without *to*.

(18)   *[He] had schame that he ne **docht** For ony paynis chang hir thocht.*
            'He had shame that he could not change her thought for any pains.'
                                                                    (*Legends of the Saints* xlvii. 173)

In example (19) from Middle Scots, *dow*, *may*, and also *can* co-occur (taken from the *OED*, s.v. *dow*, 5.).[24] Lightfoot (1979: 103) mentions this example as an instance of modal usage.

(19)   *Ye **may** not, ye **can**not, ye **dow** not want Christ*
                                                    (*OED*, Rutherford *Lett.* (1862) 1. 203, 1637)

In addition to Lightfoot (1979), several studies describe the meaning of *dow* in Scots meaning 'can'. Eitelmann (2013: 146) refers to (19) indicating that "*dugen*/*dow* 'be of value' comes to mark dynamic modality expressing an even stronger ability than *can* as late as Early Modern English". According to Moessner (1997: 113), *dought* means 'could' in the 17th century. Wilson (1923: 75–76) also refers to *dow* in Scots: *dow* and *doakht* "are used in poetry in the sense of 'can', 'could'" (e.g. *Trumlin*, *A dow noakht but glowr* "Trembling, I can do nothing but stare").

In the later period, there may have been a semantic conflict between *dow*, *may*, and *can*, which drove *dow* and *may* into disuse. Murray (1873: 215) categorizes Scots *dow* as an auxiliary, but regards it as "[n]early obsolete" (1873: 217). Purves (2002: 57) does not list *dow* in the auxiliary group,[25] and notes that "*may* is avoided in Scots" and "*can* and *coud* cover both permission and possibility". Probably *can* and *coud* drove *dow* and *may* out of the meaning area of 'ability'. The *OED* lists two examples

---

24. In addition, although not an example of Scots, Franz (1939: 178) suggests the mutual morphological influence between ModE *may* and *dow* in Shakespeare: "[n]ur einmal kommt neben *might* die Form *mought* vor, die aus me. *mahte* unter dem Einfluß von me. *douhte* (*NED.* unter *dow* S. 622) zu ae. *dugan* 'taugen' entstanden ist" [The form *mought* appears only once beside *might*, which arose from ME *mahte* under the influence of ME *douhte* (*NED* [= *New English Dictionary*] under *dow* p. 622) to OE *dugan* 'to be useful']. The example is *3Henry VI*, 5.2.45–47 (Evans 1974: 700): *That mought not be distinguish'd*; *but at last / I well might hear, delivered with a groan, / "O, farewell, Warwick!"*.

25. But Purves' (2002: 133) glossary has the entry *dowe* 'be able'. Perhaps, *dowe* is not to be regarded as a modal.

of *dow* in the infinitival form (1674; *a*1758 (s.v. 6)). These may suggest that *dow* was absent as a modal from the 17th century on, as modals usually take no infinite form (recall the second criterion described in Section 3.1.).[26] In sum, *dow* seems to have been active as a modal from the 14th to the 17th century, and later became extinct.

6. Germanic cognates of Old English *dugan*

Observed usages of some Germanic verbal cognates of OE *dugan* help us understand several of the semantic and stylistic aspects that I have presented for OE *dugan* and ME *douen*. Gothic *daug* is attested twice as a main verb meaning 'is useful'. The Old Norse (ON) cognate verb *duga* is a weak verb that does not belong to the PPV class (see Gordon 1956: 306) and is not used as a modal. Old High German (OHG) *taoc* in *Hildebrandslied* 55 *ibu dir din ellen **taoc***[27] (literally, 'if-you-your-courage-avails') is quite similar to the poetic formula in OE (see (6) in Section 4.1.). *Toug* 'is useful' occurs 11 times in Otfrid's *Evangelienbuch* (OHG), and is not only in semantic but also in rhyming relation with *mugan* 'can' (cognate to OE *magan*) with nine occurrences: *Sí thar thaz ni **dóhta**, so mir gibúrren **mohta**[28] ('if there [= in this *Evangelienbuch*] is what could not be useful, as could happen to me') (5.25.29). New High German (NHG) preserves the cognate main verb *taugen* 'to be useful' (e.g. *es taugt* 'it is useful') (< OHG *toug* and Middle High German (ᴍʜɢ) *tugen*). According to Paul (2007: 269), MHG *tugen* ceased to behave as a PPV from the 12th century, and was regularized, i.e. it became a weak verb. Old Saxon (OS) *daug* is attested five times in *Heliand*, all at the end of the B-verse in the negative, so that *daug* comprises a set phrase: *than it te uuihti ni **dôg**[29] ('then it is of benefit to nothing') (1371b: Matthew 5.13). There is no rhyming relation to ᴏs *mugan* 'can'. In Middle Low German, according to Lasch (1914: 245), *dögen* is regularized as well and did not survive as a modal.[30]

---

26. For ModE, Harris and Campbell (1995: 177–178) propose the possibility "for a modal auxiliary to develop while a homophonous modal verb continues to exist". One of their examples is the pair of *ought* ("the auxiliary reflex" of OE *agan* 'to have') and *owe*, *own* ("the verbal reflexes"). I leave the questions open to further investigation (i) if this possibility also holds of *dow* as a main verb and as a modal, and then (ii) if or how the main verb *dow* and the modal *dow* can be interrelated in their path toward obsolescence.

27. Taken from Steinmeyer (1963: 7).

28. Taken from Erdmann & Wolff (1962: 263).

29. Taken from from Behaghel & Taeger (1996: 49).

30. Cf. the Dutch regular verb *deugen* 'to be good', for which I am indebted to the editors of this volume.

Among the Germanic languages reviewed here, the OHG *Evangelienbuch* exhibits a close relation between the verbs cognate to OE *dugan* and *magan*, which may reflect their similar morphological formation at the PIE stage (see Section 2). Although OS *daug* does not collocate with OS *magan*, it is noteworthy that both OS *dugan* and OE *magan* (see Example (3)) as a main verb may occur in the same context (Matthew 5.13). OE *dugan* does not form a rhyming pair with *magan* in poetry, as does OHG *toug* with *mugan* in Otfrid, but it is in medical prose texts that OE *dugan* and *magan* are in close relation on both the syntactic and semantic levels.

Another morphological tendency that the history of *dugan* shares with its Germanic cognates is its regularization to a weak verb (cf. ON *duga* and MHG *tugen*). The regularization also occurs in ME *douen*, which is *dowes* as mentioned by Birkmann (1987: 344), but again, it seems a somewhat uncommon development for ME *douen* to be auxiliarized, while none of the Germanic cognates mentioned here is.

7.   Conclusion

This study has examined the auxiliation potential of OE *dugan*. In the OE period, this potential was latent, but it increased afterwards. It is argued that the auxiliation of *dugan* can be accounted for by similarity with OE *magan* at various levels: morphological, semantic, syntactic, and textual.

Morphologically, the similarity between *dugan* and *magan* could already be observed at the PIE stage: the ancestors of *dugan* and *magan* originate in the stative present forms, which distinguishes them from the origin of *sculan* in the perfect form. As well, OE *dugan*/ME *douen* shares features with *magan*/*mouen* as to infinitival and present-/past-participle forms. While OE *dugan* is syntactically not yet a modal, it shares the semantics of *magan* as well as of *sculan* (the latter when it combines with *wið*). In particular, *dugan* describes some positive potential for its inanimate subject of being 'good', 'useful', etc. This positive connotation is akin to the meaning of *magan* 'to be strong, efficacious' rather than *sculan* 'to owe'.

Dugan's positive, speaker-dependent connotation occurs in various textual genres. For one, Section 4 showed that *dugan*'s positive semantics occurred in some set expressions in verse and prose. The majority of *dugan* examples, however, are attested in medical writings. In medical texts, *dugan* conveying 'is good/useful' typically takes the *to*-infinitive and the preposition *wið* 'against'. The latter feature is shared with *magan* as a lexical verb. *Sculan* is also attested with *wið*, but only sporadically so. Therefore, there is strong affinitiy between *dugan and magan* at the syntactic and semantic levels. The medical manuscripts containing these characteristic tendencies of *dugan* and *magan* were handed down from the OE to the Early ME period, which contributed to the further development of *dugan*'s auxiliation.

As discussed in Section 5, *dugan*'s auxiliation in the ME period involved a reanalysis from ME *douen* with the *to*-infinitive in impersonal use + dative object to a personal construction with a nominative subject. Along with this change, ME *douen*'s notion of propriety (as evolved from positive evaluation) extended to the meaning of 'ability', leading to *dugan*'s incipient auxiliation. *Douen*'s lexical similarity to *mouen* remained valid in rhyming pairs (e.g. in *Cursor Mundi*), which assisted the syntactic change of *douen*.

In the Early Scots period, *dow* (from *douen*) appeared as a modal with the bare infinitive and remained a modal of ability until the 17th century.

Overall, OE *magan*/ME *mouen* was an influential modal precursor for the development of OE *dugan*/ME *douen*. Although the range of examples used is scarce and selective, this study has attempted to emphasize the significance of the comparison of these two PPVs in the light of their usages in the OE and ME periods.

List of abbreviations

a (e.g. *a*1225)	*ante* (as in the *OED*)	OE	Old English
c (e.g. *c*1450)	*circa* (as in the *OED*)	OHG	Old High German
ME	Middle English	ON	Old Norse
MHG	Middle High German	OS	Old Saxon
ModE	Modern English	PGmc	Proto-Germanic
ᴍs	Manuscript	PIE	Proto-Indo-European
NHG	New High German	PPV(s)	Preterite-present verb(s)

References

Dictionaries

BT = Bosworth, Joseph / Toller, Thomas N. (ed.). 1898 [Many reprints]. *An Anglo-Saxon Dictionary: Based on the Manuscript Collections of the late Joseph Bosworth*. Oxford: Clarendon Press.

BTS = Toller, Thomas N. (ed.). 1921 [Many reprints]. *An Anglo-Saxon Dictionary: Based on the Manuscript Collections of the Late Joseph Bosworth. Supplement*. Oxford: Clarendon Press.

DOE = Cameron, Angus, Amos, Ashley C., Healey, Antonette di Paolo et al. (eds). 2007. *The Dictionary of Old English: A to G Online*. Toronto: University of Toronto Press. <http://www.doe.utoronto.ca/>

DOEWC = Healey, Antonette diPaolo, Wilkin, John P. & Xiang, Xin (eds). 2009. *The Dictionary of Old English Web Corpus*. Toronto: University of Toronto Press. <http://www.doe.utoronto.ca/>

DOST = Craigie, William A. (ed.). 1931–2002. *A Dictionary of the Older Scottish Tongue*, 12 Vols. London: OUP.

LIV2 = Rix, Helmut (ed.). 2001. *Lexikon der Indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen*, 2nd edn. Wiesbaden: Dr. Ludwig Reichert.

MED = Kurath, Hans, Kuhn, Sherman M., Reidy, John & Lewis, Robert E. (eds.). 1952–2001. *Middle English Dictionary*. Ann Arbor MI: University of Michigan Press. <http://quod.lib. umich.edu/m/med/>

OED = Simpson, John A. & Weiner, Edmund S. C. (eds.). 1989. *Oxford English Dictionary*, 2nd edn. (1st edn. by James A. H. Murray et al., 1928). 20 Vols. Oxford: OUP. CD-ROM Version 4.0 (2009).

Editions

Behaghel, Otto & Taeger, Burkhard (eds). 1996. *Heliand und Genesis* [Altdeutsche Textbibliothek 4], 10th edn. Tübingen: Max Niemeyer.

Bethurum, Dorothy (ed.). 1957. *The Homilies of Wulfstan*. Oxford: Clarendon Press.

Brooks, Kenneth R. (ed.). 1961. *Andreas and the Fates of the Apostles*. Oxford: Clarendon Press.

Clubb, Merrel D. (ed.). 1925[rpt. 1972]. *Christ and Satan: An Old English Poem* [Yale Studies in English 70]. New Haven CT: Archon Books.

Cockayne, Oswald (ed.). 1865. *Leechdoms, Wortcunning, and Starcraft of Early England* [Rolls Series 35], Vol 2. London: Longman.

D'Ardenne, Simonne R. T. O. (ed.). 1961. *Þe Liflade ant te Passiun of Seinte Iuliene* [EETS, o.s. 248]. London: OUP.

Dobbie, Elliott van K. (ed.). 1953. *Beowulf and Judith* [The Anglo-Saxon Poetic Records 4]. New York NY: Columbia University Press.

Erdmann, Oskar & Wolff, Ludwig (eds). 1962. *Otfrids Evangelienbuch* [Altdeutsche Textbibliothek 49], 4th edn. Tübingen: Max Niemeyer.

Evans, Gwynne B. (ed.). 1974. *The Riverside Shakespeare*. Boston MA: Houghton Mifflin.

Fulk, Robert D., Bjork, Robert E. & Niles, John D. (eds). 2008. *Klaeber's Beowulf and the Fight at Finnsburg*, 4th edn. Toronto: University of Toronto Press.

Morris, Richard (ed.). 1874–1892 [rpt. 1961–1966]. *Cursor Mundi* (EETS, o.s. 57; 59; 62; 66; 68; 99; 101). 7 Vols. (Vol. 6: 1892). London: OUP.

von Steinmeyer, Elias (ed.). 1963. *Die kleineren althochdeutschen Sprachdenkmäler* [Deutsche Neudrucke, Reihe Texte des Mittelalters]. 2nd edn. Berlin: Weidmann.

Studies

Birkmann, Thomas. 1987. *Präteritopräsentia: Morphologische Entwicklungen einer Sonderklasse in den altgermanischen Sprachen* [Linguistische Arbeiten 188]. Tübingen: Max Niemeyer. doi:10.1515/9783111357706

Cavill, Paul. 1999. *Maxims in Old English Poetry*. Cambridge: D.S. Brewer.

Chadwick, Hector M. 1912[rpt. 1967]. *The Heroic Age*. Cambridge: CUP.

Dickins, Bruce. 1915. *Runic and Heroic Poems of the Old Teutonic Peoples*. Cambridge: CUP.

Cuyckens, Hubert, Davidse, Kristin & Vandelanotte, Lieven. 2010. Introduction. In *Subjectification, Intersubjectification and Grammaticalization*, Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens (eds), 1–26. Berlin: De Gruyter Mouton. doi:10.1515/9783110226102.0.1

Eitelmann, Matthias. 2013. Remembering (*ge*)*munan*: The rise and decline of a potential modal. In *Comparative Studies in Early Germanic Languages: With a Focus on Verbal Categories* [Studies in Language Companion Series 138], Gabriele Diewald, Leena Kahlas-Tarkka & Ilse Wischer (eds), 127–150. Amsterdam: John Benjamins. doi:10.1075/slcs.138.06eit

Fischer, Olga. 2003. The development of the modals in English: Radical versus gradual changes. In *English Modality in Context: Diachronic Perspectives* [Linguistic Insights 11], David Hart (ed.), 17–32. Bern: Peter Lang.

Franz, Wilhelm. 1939. *Die Sprache Shakespeares in Vers und Prosa*, 4th edn. Halle/Saale: Max Niemeyer.

Gneuss, Helmut & Lapidge, Michael. 2014. *Anglo-Saxon Manuscripts: A Bibliographical Handlist of Manuscripts and Manuscript Fragments Written or Owned in England up to 1100* [Toronto Anglo-Saxon Series 15]. Toronto: University of Toronto Press.

Gordon, Eric V. 1956[rpt. 1986]. *An Introduction to Old Norse*, 2nd edn. Oxford: Clarendon Press.

Harris, Alice C. & Campbell, Lyle. 1995. *Historical Syntax in Cross-linguistic Perspective* [Cambridge Studies in Linguistics 74]. Cambridge: CUP. doi:10.1017/CBO9780511620553

Hogg, Richard M. & Fulk, Robert D. 2011. *A Grammar of Old English*, Vol. 2. Chichester: Wiley-Blackwell.

Hollis, Stephanie. 2001. Scientific and medical writings. In *A Companion to Anglo-Saxon Literature* [Blackwell Companions to Literature and Culture 11], Phillip Pulsiano & Elaine Treharne (eds), 188–208. Oxford: Blackwell.

Ker, Neil R. 1957[rpt. 1990]. *Catalogue of Manuscripts Containing Anglo-Saxon*. Oxford: Clarendon Press.

Ker, William P. 1908[rpt. 1957]. *Epic and Romance: Essays on Medieval Literature*. New York NY: Dover Publications.

Krischke, Ulrike. 2013. *The Old English Complex Plant Names: A Linguistic Survey and a Catalogue* [Münchener Universitätsschriften 39]. Frankfurt: Peter Lang.

Kuteva, Tania. 2001. *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: OUP.

Kümmel, Martin. 1996. *Stativ und Passivaorist im Indoiranischen* [Historische Sprachforschung, Ergänzungsheft 39]. Göttingen: Vandenhoeck & Ruprecht.

Lasch, Agathe. 1914. *Mittelniederdeutsche Grammatik* [Sammlung Kurzer Grammatiken Germanischer Dialekte 9]. Halle: Max Niemeyer.

Lightfoot, David W. 1979. *Principles of Diachronic Syntax* [Cambridge Studies in Linguistics 23]. Cambridge: CUP

Mäkinen, Martti. 2004. Herbal recipes and recipes in herbals – Intertextuality in early English medical writing. In *Medical and Scientific Writing in Late Medieval English* [Studies in English Language], Irma Taavitsainen & Päivi Pahta (eds), 144–173. Cambridge: CUP.

Moessner, Lilo. 1997. The syntax of Older Scots. In *The Edinburgh History of the Scots Language*, Charles Jones (ed.), 112–155. Edinburgh: EUP.

Murray, James A. H. 1873. *The Dialect of the Southern Counties of Scotland*. London: Asher.

Nagle, Stephen J. & Sanders, Sara L. 1998. Downsizing the preterite-presents in Middle English. In *Advances in English Historical Linguistics (1996)*, Jacek Fisiak & Marcin Krygier (eds), 253–261. Berlin: Mouton de Gruyter.

Palmer, Frank. 2003. Modality in English: Theoretical, descriptive and typological issues. In *Modality in Contemporary English* [Topics in English Linguistics 44], Roberta Facchinetti, Manfred Krug & Frank Palmer (eds), 1–17. Berlin: Mouton de Gruyter. doi:10.1515/9783110895339.1

Paul, Hermann. 2007. *Mittelhochdeutsche Grammatik* [Sammlung Kurzer Grammatiken Germanischer Dialekte 2], revised by Thomas Klein, Hans-Joachim Solms & Klaus-Peter Wegera, 25th edn. Tübingen: Max Niemeyer.

Plank, Frans. 1984. The modals stody retold. *Studies in Language* 8(3): 305–364. doi:10.1075/sl.8.3.02pla

Purves, David. 2002. *A Scots Grammar: Scots Grammar and Usage*, revised and extended edn. Edinburgh: The Saltire Society.

Quirk, Randolph. 1955. *The Language of the Leechbook. Bald's Leechbook* (*British Museum Royal manuscript 12.D.xvii*), Appendix. In *Early English Manuscripts in Facsimile* 5, Cyril E. Wright (ed.). Copenhagen: Rodenkilde & Bagger.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language.* London: Longman.

Ringe, Don. 2006. *From Proto-Indo-European to Proto-Germanic* [A Linguistic History of English 1]. Oxford: OUP.

Traugott, Elizabeth C. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65(1): 31–55. doi:10.2307/414841

Visser, Fredericus Theodorus. 1963–1973. *An Historical Syntax of the English Language*, 3 Parts, 4 Vols. (Vol. 1: 1963; Vol. 3–1: 1969). Leiden: E. J. Brill.

Warner, Anthony R. 1993. *English Auxiliaries: Structure and History* [Cambridge Studies in Linguistics 66]. Cambridge: CUP. doi:10.1017/CBO9780511752995

Wilson, James. 1923. *The Dialect of Robert Burns as Spoken in Central Ayrshire.* London: OUP.

Sequentiality and the emergence of new constructions

That's the bottom line is (that) in American English

Reijirou Shibasaki
Meiji University

This study aims to investigate a specific construction in spoken discourse, i.e. *that's (not) the bottom line is (that)* and its variant forms, in recent American English. The function of *the bottom line is (that)* is for a speaker to introduce or anticipate upcoming talk, while *that's the bottom line* allows a speaker to summarize his/her preceding statement. These constructions give rise to a new construction when repeatedly used in the following sequence: anaphoric *that's (not) the bottom line* followed by cataphoric *the bottom line is (that)*. The newly emerging construction, i.e. *that's (not) the bottom line is (that)*, is an "apo-koinou" construction, which consists of two clauses which have a word or phrase in common. The remainder of the study then discusses how the notions of "constructional change", "constructionalization", and "constructional expansion" may be applied to the individual SN-constructions and to the resulting apo-koinou constructions.

Keywords: sequentiality, SN-constructions, apo-koinou, constructionalization

1. Introduction

This study aims to examine the construction *That's (not) the bottom line is (that)*, and its variant forms, in the past fifty years of the history of American English (as shown in Section 2.2, the earliest record of *bottom line* meaning 'the point, the crux of the argument' goes back to 1967). The construction under investigation can be seen as a merger of two constructions involving the shell noun (Schmid 2000) *the bottom line*, i.e. *the bottom line is (that)* and *that's the bottom line*; both these component structures as well as the composite structure *That's the bottom line is (that)* will, accordingly, be called "shell noun constructions" (henceforth, SN-constructions).

SN-constructions can be defined as a set of constructions comprising general/abstract nouns that serve as conceptual shells (Schmid 2000). As will be further

specified in Section 2.1, the SN-constructions under investigation here serve in discourse to express speaker stance. SN-constructions may either anticipate or summarize. Anticipating SN-constructions can be seen as a subset of "projector constructions" in interactional linguistics (e.g. Hopper 2004, 2011; Hopper & Thompson 2008; Günthner 2011; see also Günthner 2014). For example, the German projector constructions *die Sache ist/das Ding ist* 'the thing is' signal that more is to come in a subsequent stretch of talk (Günthner 2011). Other examples of projector constructions are *je veux dire* 'I want to say', *il y a* 'there is' (a presentational cleft construction; see "contact clauses" in Jespersen 1949: 132–153) and *ce qui/ce que* (a pseudocleft construction) in French (Horlacher & Pekarek Doehler 2014; Kerr 2014), and *are* 'that' (the distal demonstrative) in Japanese (Hayashi 2004) (see Günthner 2011: 19). Like other projector constructions, *the bottom line is (that)* foreshadows upcoming talk, which is here expressed in the complement clause introduced by (*that*); an example is (1), cited in the lexical entry of the shell noun *bottom line* in *Oxford Advanced Learner's Dictionary* (*OALD*).

(1)   The *bottom line is that* we have to make a decision today.

(*OALD* 2010, s.v. *bottom line* 1)

The other SN-construction making up *That's the bottom line is (that)* is *that's the bottom line*. An example is (2), taken from the lexical entry *bottom line* in *The Longman Dictionary of Contemporary English* (*LDOCE*):

(2)   In radio, you have to keep the listener listening. *That's the bottom line.*

(*LDOCE* 2012, s.v. *bottom line* 1)

In (2), *that's the bottom line* does not serve as a projector; rather, it summarizes preceding information.

It appears then that, as talk unfolds in spoken real-time discourse, different SN-construction may fulfill different functions: in particular, *the bottom line is that* cataphorically projects information, while *that's the bottom line* has an anaphoric function summarizing information. Importantly, the summarizing and projecting function can occur in sequence (but not in the reverse order; see Sections 3.2 and 3.3), as in (3).[1]

(3)   Dr. LEVY: Well, we certainly do believe that people are born with a predisposition to become addicted, but *that's not the bottom line. The bottom line is* whether Timmy continues to use or not.

(*Corpus of Contemporary American English* (*COCA*);
SPOK, *CNN King*, 1990)

---

1.   Note that little has been said about the sequential use of the summarizing and projecting constructions in the literature on shell nouns (Schmid 2000; Flowerdew & Forest 2014).

Such sequential use is also witnessed in constructions with *fact*, *difference*, *issue*, *point*, *problem*, *question*, *thing*, *trouble*, *truth*, etc., as follows.

(4) Ms-WALDEN: I want the jury and the audience to know that I love Michael Jackson as much as everybody else does. But *that isn't the issue. The issue is that* we have a child, a child who has accused or made allegations that they have been sexually abused by a 35-year-old man.

<div align="right">(COCA; SPOK, Ind Geraldo, 1993)</div>

(5) ALLISON: (nods) But *that's not the problem. The problem is that* he blames himself. (COCA; FIC, Mov.Case Curiosities, 2000)

The sequential SN-constructions involving *the bottom line* can also be merged into one single form, called an "apo-koinou" construction, as in (6):

(6) VELEZ MITCHELL: I think *that's the bottom line is that* our culture is not designed to induce recycling. (COCA; SPOK, CNN Cooper, 2012)

According to *The Oxford English Dictionary* (*OED*, s.v. *apo koinou, apo-koinou*, Gr. *ἀπὸ κοινοῦ* 'in common'), apo-koinou constructions "[consist] of two clauses which have a word or phrase in common" (see Corminboeuf 2012: 215–216 for related terms). A variety of such apo-koinou constructions have been discussed in philological studies (e.g. Kellner 1892; Jespersen 1949; Visser 1963; Meritt 1967; Blockley 2001; Ukaji 2003; Miura 2009; Yaguchi 2010 and references therein) as well as in functionally oriented studies (e.g. Lambrecht 1988, 2001; Hopper 2007; Corminboeuf 2012; Horlacher & Pekarek Doehler 2014) in a range of languages.[2] However, the constructions that can be schematized as "that's the X is (that)", as exemplified in (6), had gone unnoticed until Ross-Hagebaum (2005), which seems to be the only in-depth study thus far that has carried out a careful examination of these constructions (but note that this study has also left unaddressed the sequential use in real-time discourse of the component SN-structures – *that's the bottom line* and *the bottom line is (that)*, as in (3)). Building on this preceding research, this study will probe deeper into the sequential nature of the SN-constructions involving *the bottom line* in discourse and the emergence of a newly derived apo-koinou construction in recent American English.

---

2. Here are examples of apo-koinou constructions in Middle and Present-day English.

(i) *And hedde hem under an holw hok was an huge denne.*
'And they hid under a hollow oak was a huge den.'

<div align="right">(14c., William of Palerne, 1793; Meritt 1967: 64)</div>

(ii) JULIE: *and I said,*
this is what I need is a gay stallion.

<div align="right">(Santa Barbara Corpus of Spoken American English (SBCSAE) 056)</div>

This study is structured as follows. Section 2 introduces the necessary background for the study of *the bottom line is (that)* and *that's the bottom line*. Section 3 presents the results from corpus surveys, while Section 4 discusses the emergence of apo-koinou constructions. Section 5 concludes this study.

2. Theoretical and historical background

2.1 Shell noun constructions

Schmid (2000) is a systematic study of SN-constructions in Present-day British English based on the *Bank of English* corpus; in particular, it considers the ways each SN-construction is put to use in discourse. SN-constructions are classified into various types, as shown in Table 1. In this table, shell noun phrases are in boldface, while their concrete contents, called "shell contents" (Schmid 2000: 7), are underlined.

Table 1. Shell noun constructions (based on Schmid 2000: 22)[3]

Function	Pattern	Abbreviation	Example
Cataphoric	Shell noun + postnominal clause	N+cl	Mr. Bush said Iraq's leaders had to face **the fact** that the rest of the world was against them.
	Variants: *that*-clause, *to* infinitive-clause, *wh*-clause	N-*that* N-*to* N-*wh*	
	Shell noun + *be* + complementing clause	N-*be*-cl	**The advantage** is that there is a huge audience that can hear other things you may have to say.
	Variants: *that*-clause, *to* infinitive-clause, *wh*-clause	N-*be-that* N-*be-to* N-*be-wh*	
Anaphoric	Referring item + (modifier) + shell noun	*th*-N	(Mr. Ash was in the clearest possible terms labelling my clients as anti-semitic.) I hope it
	Referring item as subject + *be* + shell noun (phrase)	*th-be*-N	is unnecesssary to say that **this accusation** is also completely unjustified. (I won the freshmen's cross-country. – Mm.) That was **a great achievement** wasn't it?

---

3. Note that Schmid (2000) uses the term "pattern" rather than "construction".

The nouns that go under the label of shell nouns (*fact*, *advantage*, etc.) are not necessarily specific or informative in themselves, but they serve "the textual function of *linking* these nominal concepts with clauses or other pieces of text which contain the actual details of information, thereby instructing the hearer to interpret different sections of a text together" (Schmid 2000: 14, emphasis in original). Among the 670 shell nouns examined in Schmid (2000), the ones that most frequently occur in the "N-*be-that*" construction are *problem*, *thing*, *truth*, *fact*, *trouble*, *point*, etc. (Schmid 2000: 59 for details).

Let us consider some examples of SN-constructions comprising the shell noun *bottom line*.

(7)   ROBELOT: But you know what? *The bottom line is*, Pat Riley's just the greatest coach that ever – other than John Wooden – that ever coached the game of basketball.
McEWEN: Yeah, I don't know about that. But *the bottom line* is that the Chicago Bulls got something for whoever won that series.
(*COCA*; SPOK, *CBS Morning*, 1997)

(8)   QUAYLE: … I'm saying to you that there'll be on the positive side 170,000 net new jobs. And the question- *the bottom line is*, is this going to create more jobs and opportunities for Americans or is it going to create less jobs and opportunities? (*COCA*; SPOK, *ABC Brinkley*, 1992)

(9)   Mr. VAN NATTA: I don't think a politician has ever said that, because *that is- that's the truth. That's the bottom line* and that's just not said in Miami or in Florida. (*COCA*; SPOK, *NPR Weekend*, 1994)

Example (7) includes the SN-construction "N-*be-that*" in Table 1 and its variant "N-*be-Ø*". The first speaker, ROBELOT, uses *the bottom line is that* to emphasize that Pat Riley is just the greatest coach; the second speaker McEWEN, while replying to ROBELOT, also uses *the bottom line is* to introduce additional information about the Chicago Bulls. Example (8) is also an instance of the SN-construction "N-*be-that*" in Table 1; however, this sequential use is different from the patterns in the table because the construction introduces a direct question instead of an indirect question. Here, the speaker starts off the interrogative SN-construction with the shell noun *the question*, which is then immediately rephrased as *the bottom line is*. In (9), which instantiates the anaphoric structure "*th-be-*N", the speaker first strongly voices his opinion by using *that's the truth*. This statement is immediately followed by *that's the bottom line*. Such repetition or rephrasing of SN-constructions as in (8) and (9), which is not discussed in Schmid (2000), can be witnessed in the data (see note 11).

As was pointed out before, SN-constructions specialize in either anticipating upcoming talk by the same speaker or summarizing the speaker's preceding statement. In that capacity, the choice of a particular shell noun in its SN-construction reflects the speaker's interpretation of the immediate context in the unfolding discourse. In a nutshell, SN-constructions are argued to be discourse-based constructions that reflect speaker stance or introduce interrogative clauses, as the discourse unfolds (they fulfill a function similar to "prefabs" in Erman & Warren 2000: 31, cited in Los 2015: 7; see also Bybee 2010: 34–36).

The syntactic patterns in Table 1 (Schmid 2000) pay no attention to the blending of the peripheries of utterance as seen in a longer stretch of discourse, i.e. from *that's the bottom line* plus *the bottom line is (that)*, as in (3), to *that's the bottom line is (that)*, as in (4); related works such as Flowerdew & Forest (2014) do not discuss the issue either. Actually, SN-constructions involving *bottom line* are not addressed in either Schmid (2000) or Flowerdew & Forest (2014). Therefore, this study will reexamine the discourse functions of these SN-constructions.

2.2   A brief history of *bottom line*

The *OED* defines *bottom line* as follows: "orig. U.S., the last line of a profit-and-loss account, showing the final profit (or loss); fig., the final analysis or determining factor; the point, the crux of the argument". The last meaning 'the point, the crux of the argument', which is figuratively derived from the semantics of profit and loss accounting, is attested in the SN-constructions *the bottom line is (that)* and *that's the bottom line*. The *Online Etymology Dictionary* tells us that this figurative meaning dates from 1967, while the *OED* provides us with the following earliest examples of each construction.

(10)   George Murphy and Ronald Reagan certainly qualified because they have gotten elected. I think *that's the bottom line*.
(*OED*, s.v. *bottom* III *attrib.* and *Comb.* 20. Special comb, 1967.)

(11)   *The bottom line is that* invention is much more like falling off a log than like sawing one in two. (*OED*, ibid., 1982)

Making use of this information, the present study conducts a detailed survey of the uses of the two SN-constructions involving *bottom line*. Our study is mainly based on spoken data (documenting interactional and conversational usage) from *The Corpus of Contemporary American English 1990–2012* (*COCA*) and *The Corpus of*

Historical American English 1810–2009 (*COHA*).[4] Note that while the sequential use of the SN-constructions can be found in *COHA*, albeit only with five types of shell nouns, no apo-koinou constructions with shell nouns can be witnessed in the corpus.

3. Survey results

3.1 The distributional patterns

Tables 2 and 3 provide an overview, with attendant frequencies, of the sequential use of the SN-constructions instantiating the types "*th-be*-N" and "N-*be-that*", as exemplified in (3), and of the amalgamated constructions instantiating "*That-be*-N-*is*-(*that*)", as in (4). Table 2 presents the data from *COCA*, while Table 3 summarizes the data from *COHA*; the numbers (= n) stand for the raw frequency of the sequential use of each SN-construction.

I conducted a comprehensive survey of the sequence of *That's (not) the X. The X is (that)* and *That is (not) the X. The X is (that)* in the corpora and found that this specific sequential use of the SN-constructions recently developed with a limited set of shell nouns, as seen in the tables. Notice that the tables also include the constructional sequence *It's not X. The X is*, which serves similar syntactic and semantic functions to *That's (not) the X. The X is (that)* and *That is (not) the X. The X is (that)*.[5] In Table 3, the SN constructions with the verbal form *ain't* and with *that is beside/beyond the point* as well as *that is not the point* are added; the number of *that is beside/beyond the point* is in parentheses.

The survey results tell us that, although 670 types of SNs are examined in Schmid (2000), only a limited set of SNs occur in the sequentially used SN-constructions (see (3)), and an even lower number of types and tokens of corresponding apo-koinou SN-constructions are attested. In the next subsections, we will discuss the sequential use and the flow of information of these SN-constructions.

---

4. For historical studies of *the thing is (that)* and *the point is (that)*, see Shibasaki (2014a, 2014b).

5. Some other examples of similar sequential uses, e.g. *Here's the point. The point is*, *Wh-cleft is not the point. The point is*, and *Money is not the issue. The issue is*, are not included in Tables 2 and 3, because these examples are syntactically or semantically different from the patterns in Table 1: *here* is cataphoric, while *wh-cleft* and lexical nouns are outside the scope of the present study and they are very infrequent.

Table 2. Sequential and apo-koinou uses of SN constructions in *COCA* (accessed 25 December, 2014)

SN types	Sequential use (*n* = raw frequency) *th-be*-N + N-*be-that*	Total
point	That's/is the point. The point is (*n* = 7) That's <u>not</u> the point. The point is (*n* = 42) That is <u>not</u> the point. The point is (*n* = 6) This is <u>not</u> the point. The point is (*n* = 1) That s <u>not</u> the point. The point is (*n* = 1)	57
issue	That's the issue. The issue is (*n* = 2) That's <u>not</u> the issue. The issue is (*n* = 29) It's <u>not</u> the issue. The issue is (*n* = 1)	32
problem	That's the problem. The problem is (*n* = 4) That's <u>not</u> the problem. The problem is (*n* = 13)	17
question	That's the question. The question is (*n* = 2) This is the question. The question is (*n* = 1) That's <u>not</u> the question. The question is (*n* = 10) That is <u>not</u> the question. The question is (*n* = 3)	16
truth	It's <u>not</u> the truth. The truth is (*n* = 1) That shit ai<u>n't</u> the truth. The truth is (*n* = 1) That's really simplifying and <u>not</u> exactly the truth. The truth is (*n* = 1)	3
thing	CLAUSE, is the thing. The thing is (*n* = 1)[7]	1
fact	The first is the fact. The fact is (*n* = 1)	1
trouble	That's the trouble. The trouble is (*n* = 1)	1
bottom line	That's <u>not</u> the bottom line. The bottom line is (*n* = 1)	1
difference	N/A	1
reality	That's the reality. The reality is (*n* = 1)	1

---

6. The expression *but Ø is the point is* means that in the sequence, *that* or *this* is unexpressed in the data:

(i) Mr. ENGST: The- as I said, in general, the best way to do it, really, is through one of the various books. And those are generally about $30. Once you're-
ZWERDLING: -OK. Well so- but *is the point is*, we're talking about a lot of money here…   (*COCA*; SPOK, *NPR ATC*, 1994)

7. The expression *CLAUSE, is the thing. The thing is* indicates that after the independent clause, *is the thing* appears without any relative or demonstrative pronouns expressed, as in *I don't talk religion all that much, <u>is the thing</u>. <u>The thing is</u>, most people I talk to don't talk religion* (*COCA*; FIC, *Some thrills*, 2000).

Apo-koinou (n = raw frequency)	Total
that's the point is ($n = 3$) but Ø is the point is ($n = 1$)[6]	4
That's the issue is ($n = 2$)	2
That's the problem is ($n = 8$)	8
That's <u>not</u> the question is ($n = 1$) It's the question is ($n = 1$)	2
That's the truth is ($n = 1$) It's the truth is ($n = 1$)	2
That's the thing is ($n = 8$)	8
N/A	0
That's the trouble is ($n = 1$)	1
That's the bottom line is ($n = 1$) That's the bottom line, is… ($n = 1$) That the bottom line is ($n = 1$)	3
That's the difference is ($n = 1$)	1
N/A	0

Table 3. Sequential and apo-koinou uses of SN constructions in COHA (accessed 26 May 2016)

Periods	SN types — Sequential use: *That's (not) the point. The point is*	Apo-koinou: *That's the point is (that)*	Sequential use: *That's (not) the problem. The problem is*	Apo-koinou: *That's the problem is (that)*	Sequential use: *That's (not) the question. The question is*	Apo-koinou: *That's the question is (that)*	Sequential use: *It's (not) the truth. The truth is*	Apo-koinou: *That's the truth is (that)*	Sequential use: *That's (not) the thing. The thing is*	Apo-koinou: *That's the thing is (that)*
1840s					1					
1850s										
1860s	1									
1870s					1					
1880s										
1890s										
1900s										
1910s	3									
1920s	4									
1930s	6 (1)				1					
1940s	2 (1)									
1950s	6 (1)				2				1	
1960s	1 (1)									
1970s	3 (2)		1							
1980s	8		1							
1990s	3		1				1			
2000s	2		3		1					
Total	**39**	**0**	**6**	**0**	**6**	**0**	**1**	**0**	**1**	**0**

3.2 Sequential use and the flow of information

As shown in Table 2, *COCA* contains one instance of the sequential use of *that's the bottom line* and *the bottom line is (that)*; it is repeated here as (12).

(12) Dr. LEVY: Well, we certainly do believe that people are born with a predisposition to become addicted, but *that's not the bottom line. The bottom line is* whether Timmy continues to use or not.　　(*COCA*; SPOK, *CNN King*, 1990)

In (12), the speaker expresses the general view that people are born with a predisposition to become addicted, but he adds right away that this is not the main point, as conveyed by *that's not the bottom line*. The speaker then introduces his own point of view with *the bottom line is*. The former construction thus points backward, while the latter points forward. In particular, the speaker blocks the flow of information and summarizes the preceding information by *that's not the bottom line*; at the same time, this statement has pragmatic value in that it signals to the hearer that relevant information is upcoming. This, then, enables the speaker to put forward his/her opinion introduced by *the bottom line is*.

　　Similar discourse functions can be found in the sequential use of other SN-constructions, as shown in (13) and (14).

(13) Our guys wear Nike and Pony shoes and Air-Max helmets, just like Notre Dame or Michigan. *That's not the problem. The problem is*, we don't have any Tyrone Wheatleys.　　　　　　　　　　(*COCA*; NEWS, *New York Times*, 1994)

(14) DONNA BRAZILE-1-A# (Off-camera) So how many jobs were created during its period when we had the Bush tax cuts, from 2001 to 2008? How many jobs? STEPHEN HAYES-1WE# I don't – I don't know the number.
DONNA BRAZILE-1-A# (Off-camera) Hardly any. We lost jobs. *That's the point. The point is*, we should have tax cuts that help spur economic growth. *Bottom line*, it's about jobs, jobs, jobs.　(*COCA*: SPOK, *ABC This Week*, 2010)

In (13), as in (12), the first construction of the sequence, i.e. *that's not the problem*, contains the negative adverb *not*; hence, the flow of information is blocked, after which the speaker expresses his/her own opinion. In contrast, *that's the point* in (14) does not include *not*. The information that this SN-construction refers back to, namely "we lost jobs" (in contrast to a promised increase), is partly maintained in the subsequent construction *the point is*, which serves to raise a related idea, namely "we should have tax cuts that help spur economic growth".[8] Syntactically

---

8.　Note in passing that as in (14), *bottom line* can be used as a projector construction without determiner, copula and complementizer.

and semantically, *that's the point* makes anaphoric reference to the preceding discourse (see Table 1); at the same time, it serves a pragmatic function similar to that in (13), namely to signal relevant upcoming information. It would appear, then, that the *th-be*-N plus N-*be-that* sequence is cataphorically oriented, regardless of whether the first component *th-be*-N contains *not* or not.

3.3 The emergence of apo-koinou constructions – loss of the negative adverb *not*[9]

The "*th-be*-N" construction in Table 2 appears to be used frequently with the negative adverb *not*, especially in the four most frequent types involving the shell nouns *point*, *issue*, *problem*, and *question*. This tendency for *not* to occur in the "*th-be*-N" construction can be explained by the fact that it enables the speaker to hold the floor and delay the subsequent delivery of the message that for him/her is crucial in which he/she contrasts his/her opinion with views uttered in the preceding context. Because of this discourse strategy, once interlocutors hear the "*th-be*-N" construction containing *not*, they can easily prepare for the following statement. In this regard, the "that's not the X" construction pattern also facilitates speaker–hearer interaction in discourse (as does the "that's the X" pattern in (14), for that matter).

The three examples of the apo-koinou construction with *bottom line* do not include the negation marker; an example is (4), repeated here as (15).

(15) UNIDENTIFIED FEMALE: But there aren't enough garbage receptacles on the street to say this is the plastic, put it in here.
VELEZ MITCHELL: I think *that's the bottom line is that* our culture is not designed to induce recycling. (*COCA*; SPOK, *CNN Cooper*, 2012)

(16) BASTINE: Well, I had a lobbyist for one of the drug manufacturers tell me that.
WALLACE: Say-
BASTINE: *That the bottom line is* it hits them in the pocketbook.
(*COCA*; SPOK, *ABC Primetime*, 1993)[10]

(17) Ms. ROBERTS: I think they really don't think they have a choice. I think *that's the bottom line, is that* they're not crazy about putting all their eggs in one basket, but they don't really have any other basket to put them in.
(*COCA*; SPOK, *ABC Brinkley*, 1993)

---

9. Part of the analysis in Section 3.3 owes to my discussion with Elizabeth C. Traugott (on 8 December, 2014).

10. Note that in (16), the copula is unexpressed (or not transcribed) as in *that the bottom line is*.

In all these examples, the apo-koinou construction does not refer back to previously mentioned utterances but refers forward to subsequent utterances. Although the cataphoric and anaphoric component SN-constructions are merged into one, the apo-koinou construction itself turns out to be cataphorically oriented. The loss of *not* in almost all cases of the apo-koinou constructions in Table 2 gives/is evidence in support of this interpretation.[11, 12]

The cataphoric reference of the apo-koinou construction can clearly be seen in the following example.

(18)   MR SOLMAN: Okay. So talk about some. Talk about them specifically, if you wouldn't mind, with regard to trade. I mean, what do you do?
MR GARTEN: *That's the problem*. You see, *that's the problem is that*, if you want to, if you ask me, what should we do in trade, it would be more of it, it would be promote exports even more aggressively than we do.
(*COCA*; SPOK, *PBS Newshour*, 1996)

In this exchange, the utterances after *that's the problem is that* present new information; instead of the apo-koinou construction, therefore, the speaker might equally well have chosen the cataphoric SN-construction *the problem is (that)* (see footnote 12 in this regard). This particular example thus confirms that apo-koinou constructions, schematically represented as "that's the X is (that)" are oriented toward upcoming information, i.e. are cataphoric in nature.

---

11. Given that the absence of *not* is seen as evidence in favor of a cataphoric interpretation of the apa-koinou construction, it cannot be excluded that some sequential SN-constructions that do not have *not* in the first segment, i.e. *th-be*-N, may also be construed as cataphoric. (p.c., Hubert Cuyckens, 29 April, 2016).

12. In Table 2, the apo-koinou construction with *question* appears to be an exception because it contains *not*. Consider example (i) in this respect. In (i), the first use of *that's not the question* may have triggered the production of the subsequent use as an apo-koinou construction. This consecutive use of similar syntactic structures across utterances or across interlocutors is often witnessed in naturally occurring discourse, as investigated in details in Du Bois (2007, 2014) under the name of "dialogic syntax".

(i)   PRES BUSH: The best thing for peace is to move – to move the process forward is just to have this deferral. But I'm going to fight for what I believe and it may be popular politically but probably it's not. But *that's not the question* here. *That's not the question is* whether it's good 1992 politics. What's important here is that we give this process a chance. (*COCA*; SPOK, *PBS Newshour*, 1991)

4. Discussion

4.1 Apo-koinou constructions as a case of constructionalization

In this section, I will consider apo-koinou constructions as well as their component SN-constructions from the perspective of "constructionalization" and "constructional change", as outlined by Traugott and Trousdale (2013) and Traugott (2014). According to Traugott (2014), constructionalization and constructional change are defined as follows.

(19) Constructionalization
is the development of form$_{new}$ – meaning$_{new}$ pairs, i.e. constructions.
(Traugott 2014: 89)

(20) Constructional changes
are changes to features of constructions, such as semantics (e.g. *wif* 'woman' > 'married woman') or morphophonology (e.g. *had* > *'d*). Such changes precede or follow constructionalization. (Traugott 2014: 89)

As was seen in Section 3, apo-koinou constructions can be said to have emerged from the sequential use of two independent SN-constructions, namely the projecting construction "X is (that)" and the summarizing construction "that's the X". Each construction has its own discourse-pragmatic meaning: the projecting construction anticipates upcoming information in the subsequent discourse, while the summarizing construction encapsulates the immediately preceding discourse content. Once these two constructions are fused into one single construction, as in (15), the distinct referentiality of each of the original constructions, i.e. cataphoricity and anaphoricity, is given up and converges into cataphoric reference (see Section 3.3). This fused construction, i.e. "that's the X is (that)", is semantically different from its component constructions, i.e. "the X is (that)" and "that's the X"; moreover, the newly formed construction is formally different from its original constructions. In fact, as shown in Table 3, the apo-koinou constructions are much more recent than their component SN-constructions. Given that a new form–meaning pair is established, apo-koinou constructions can be regarded as a case of *constructionalization* (19).

In addition, the projecting and summarizing SN-constructions, which contribute to the apo-koinou construction, may be subject to constructional change. Let us take a look at the examples in (21)–(23).

(21) K. DOLAN: I think *the bottom line – my bottom line is* just be awfully careful what gets prescribed. (*COCA*; SPOK, *CNN Dolans*, 2005)

(22)   DONALD RIEGLE (Michigan): So – so you – you attribute this to just the
       normal – sort of grumbling and griping that goes on within an organization.
       Is that – that's your bottom line on this?
       Mr CASEY: *That's my bottom line.*   (*COCA*; SPOK, *CBS Street*, 1992)

(23)   RANDALL: … What's the bottom line here?
       KENNETH GROSS, CNN ELECTION LAW ANALYST: *The bottom line is, is
       that* the recounting will stop, and that's the critical juncture here for the Gore
       campaign.   (*COCA*; SPOK, *CNN Live Sat*, 2000)

In (21), the speaker starts the projecting SN-construction with *the bottom line*
but instantly rewords this as *my bottom line* replacing *the* by the possessive *my*
(this SN-construction then serves to introduce an imperative construction). *My
bottom line*, which is a constructional variant of *the bottom line*, may also have a
summarizing function, as in (22). It is likely that since the first speaker DONALD
RIEGLE uses the construction with the possessive *your* in *that's **your** bottom line
on this?* to enquire about his interlocutor's stance in relation to the topic of the
discussion, the second speaker CASEY uses *that's **my** bottom line*, also comprising
a possessive.

Discourse fragment (23) exemplifies the so-called "double *is*" (or "reduplicative
copula" or "two-'be' (2-B)'") construction (see Bolinger 1987; Tuggy 1996; Massam
1999; Shapiro & Haley 2002; Andersen 2002; Coppock et al. 2006; Curzan 2012 and
references therein). Tuggy (1996: 728–729) states that the presence of a solecisti-
cally redundant copula indicates the speaker's stuttering, hesitation, mishearing,
etc., depending on the context; in (23), the reduplicated copula introduces a pause
to accentuate the utterance that immediately follows. Furthermore, the number of
double *is*-constructions in general are extremely infrequent in comparison to their
corresponding single *is*-constructions, according to Curzan (2012).

Tables 4 and 5 show us how variant forms of *the bottom line is (that)* and
that's the bottom line have developed in the history of American English. It is
obvious that *the bottom line is that/Ø* and *that's the bottom line* are older than
their corresponding variant constructions *my/your bottom line* is and *that's my
bottom line*, respectively, as in (21) and (22). In fact, such possessive examples
are not found in *COHA*. Note in passing that twelve examples of the sequential
use of *my bottom line is* with a finite clause are found in *COCA* (accessed 11 June,
2016), which implies that "POSS bottom line is (that)" is of a later date than *the
bottom line is (that)*. The same goes for *that's the bottom line* and *that's my bottom
line* in Table 5.

Table 4. Constructional variations of *the bottom line is (that)* in *COHA* (accessed 11 June, 2016)

	1960s	1970s	1980s	1990s	2000s	Total
The bottom line is that	0	2	14	21	23	60
The bottom line is Ø	0	1	10	24	22	57
Ø bottom line is that	0	0	0	0	0	0
Ø bottom line is Ø	0	0	0	0	0	0
POSS bottom line is that	0	0	0	0	0	0
Total	**0**	**3**	**24**	**45**	**45**	**117**

Table 5. Constructional variations of *that's the bottom line* in *COHA* (accessed 7 July, 2014)

	1960s	1970s	1980s	1990s	2000s	Total
That's the bottom line	0	1	3	6	5	15
That's Ø bottom line	0	0	0	0	0	0
That's POSS bottom line	0	0	0	0	0	0
Total	**0**	**1**	**3**	**6**	**5**	**15**

In addition to reduplication, accentuation in an SN-construction may also be achieved by emphatic delivery, as in (24), conveyed by *thee* [ði:] (instead of *the*).

(24)   (CROSSTALK) HOSTIN: And I think that argument makes absolutely no sense, I have got to tell you. *Thee bottom line is*, she had an affair with John Edwards.
(*COCA*; SPOK, *Fox Hannity*, 2009)

In (24), the speaker emphatically delivers *the* as *thee* [ði:], allowing him/her to build in a short pause before expressing what is actually at stake, namely that "she had an affair with John Edwards"; note that the transcript includes a comma after the construction.

It would appear that the SN-constructions in (21)–(24) can be seen as cases of constructional change (rather than constructionalization). The variants in (21), (22), and (23) of the SN-construction *the bottom line is (that)* are obviously formally different, but retain their projecting/cataphoric function; the formal variant in (22) of the SN-construction *that's the bottom line* retains its summarizing/anaphoric function. Putting all this together, it seems to be impossible to say at what point along the "constructional change – constructionalization" continuum a case of constructional change becomes a clear case of constructionalization. Therefore, this study takes only the apo-koinou construction as a case of constructionalization, while considering all other cases discussed here as cases of constructional change positioned at various points along a cline towards constructionalization.

For example, the double *is*-construction seems to be semantically and pragmatically specialized in or skewed towards expressing hesitation, and no longer simply a projecting construction, as follows.

(25)   BARACK OBAMA: I think, John, *the fact is, is that* I opposed this war from the start.   (*COCA*; SPOK, *NBC Today,* 2007; Shibasaki 2015b: 164)

(26)   TERRY TOLLIVER: Right, but still, you're, you know, *the fact is is that*, as I said, legislation can only go so far…
(*COCA*; SPOK, *NPR Weekend,* 1995; Shibasaki 2015b: 164)

In these examples, the double *is*-construction *the fact is is that* is used with other hedges or discourse/pragmatic markers, e.g. *I think*, *you know*, *as I said*, all of which show signs of the speaker's hesitation. In (25), for example, Barack Obama, by using *the fact is, is that*, hesitantly discloses his real intention, i.e. *I opposed this war from the start*; the occurrence of a comma between the copulas (i.e. *is, is*) may reflect his diffident manner of speaking. The sequential use of *Right, but still, you're, you know, the fact is is that, as I said* in (26) clearly shows the speaker's hesitation in speech.

4.2   Constructional expansion

In Section 3, a range of SN-constructions were described, with a focus on *bottom line*. It was pointed out that their sequentiality (a summarizing SN-construction followed by a prospective one) is crucial for the emergence of apo-koinou constructions. In this section, I will consider the importance of analogy in expanding the range of constructions.

According to Croft (2001), constructions can be characterized as symbolic form–meaning pairings; (parts of) constructions may make up a network whose elements are taxonomically related. This constructional view holds true for the apo-koinou constructions discussed. As shown in Table 2, nine types of shell nouns can be used in apo-koinou constructions, at varying frequencies of use. Apo-koinou constructions involving the shell nouns *reality* and *fact*, however, have not been attested in the corpus, while the sequential use of *that's the reality/fact* and *the reality/fact is* has.

Figure 1 shows a constructional taxonomy of "that's the X is (that)" based on the frequency data of our corpus. The dashed lines stand for potentially new constructions; "that's the X^{+1} is (that)" represents a prospective apo-koinou construction with X^{+1}.

Arguably, even fairly recent constructions at low levels of schematicity (such as *that's the thing is (that)* and *that's the problem is (that)* can form a more schematic construction "that's the X is (that)", which in turn may sanction the development

of new constructions at a lower level of schematicity, thus gradually giving rise to a cluster of constructions as in Figure 1. It is likely that frequency plays an important role in this constructional expansion as a more frequent construction (e.g. *that's the thing is (that)*) may accelerate the use of a less frequent one (e.g. *that's the difference is (that)*) and may potentiate new constructional candidates (e.g. *that's the reality is (that)*, *that's the X⁺¹ is (that)*). In sum, once a network of apo-koinou constructions involving shell nouns has been established, however emergent, relatively new shell nouns such as *bottom line* will also give rise to apo-koinou constructions.

that's the *thing* is (that)

that's the *problem* is (that)

that's the *point* is (that)

that's the *truth* is (that)

that's the *bottom line* is (that)

that's the X is (that)   that's the *issue* is (that)

that's not the *question* is (that)

that's the *trouble* is (that)

that's the *difference* is (that)

(that's the *fact* is (that))

(that's the *reality* is (that))

(that's the X⁺¹ is (that))

more frequent

less frequent
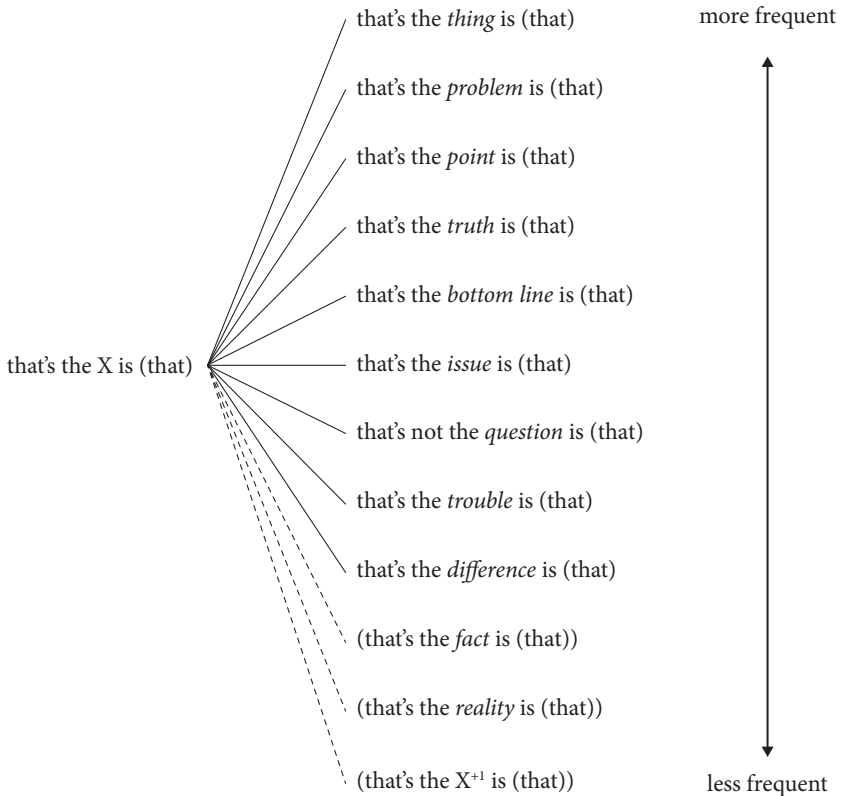
Figure 1. A constructional taxonomy of "that's the X is (that)"

It is well known that repeated occurrences of a given expression play a determining role in language change (e.g. Bybee 2010: 34, see Section 2.1). As shown in Table 2, however, the raw frequency of each apo-koinou construction is extremely low and even the number of sequentially used examples appears to be small. One might thus question whether the constructional expansion of low-frequency items, as schematized in the taxonomy in Figure 1, is as common as seen in the case of

high-frequency items. Nevertheless, it is pointed out, for instance by Hoffmann (2005: 140–165), that low-frequency items have the potential to gradually increase in number by their analogy with structurally similar high-frequency items. Table 3 illustrates some randomly sampled examples of low-frequency complex prepositions – *in X of* – from the written component of the BNC (Hoffmann 2005: 142).

Table 6. Low-frequency examples of complex prepositions in BNC

in + X + of (Hoffmann 2005: 142)

Examples	Raw frequency
in sight of	54
in ignorance of	46
in celebration of	43
in discussion of	26
in quest of	18
in want of	6
in proof of	5
in presence of	5

What Hoffmann's (2005) data reveal is that low-frequency items seem to form a taxonomic pattern presumably from more to less frequent constructions, which can increase its capacity on its own by creating further sub-constructions sanctioned by the schematic *in X of*. Note in this respect that a definition of high-frequency is not absolute but fairly arbitrary. In fact, Bybee (2015: 40) clearly states that "it is not currently known exactly how to determine what is low and what is high" in frequency. All the same, the expansion of relatively low frequency constructions shown in Figure 1 can serve as corroborating evidence for the language change expressed in Hoffmann (2005) and Bybee (2010, 2015).[13]

The apo-koinou construction "that's the X is (that)" can be regarded as an emergent phenomenon because actual instances do not go back a long way; as pointed out in Section 2.2, the shell noun *bottom line* can only be traced back to 1967. One cannot be quite sure whether the present variants shown in Table 2 continue to trigger further constructional variations. Nevertheless, it is likely that language users will continue to create and recreate grammar, as shown in Figure 1, when they gain an understanding of the construction in a wider taxonomy and comprehend its relation to other constructions at various levels of schematicity, whereby *that's*

---

13. As Wray (2009: 48) points out, "humans will strive to retain the creative edge"; this, I believe, can be reflected in speakers coining brand new infrequent constructions such as the apo-koinou constructions involving *bottom line*.

the bottom line is (that) is a construction that inherits properties from higher up in the taxonomy, i.e. "that's the X is (that)".[14]

5.   Concluding remarks and further issues

In this study, the following aspects of the apo-koinou construction "that's the X is (that)" have been examined with special reference to *that's the bottom line is (that)*. Firstly, on the basis of the corpora consulted, the apo-koinou construction turns out to be a recent phenomenon. Secondly, a sequential alignment of two independent SN-constructions, namely the anaphoric "that's the X" and the cataphoric "the X is (that)", can be considered a necessary condition for the emergence of "that's the X is (that)". Once merged, the apo-koinou construction "that's the X is (that)" becomes cataphorically oriented whilst sharply decreasing the anaphoricity of its component construction "that's the X". Furthermore, the formation of the apo-koinou construction involving *bottom line* is facilitated by its analogy with more frequent such constructions (involving, for instance, the shell nouns *problem*, *thing*). Finally, our results provide corroborating evidence for the view that constructions are formed and progress in a network, not in isolation.

It is also worth pointing out that, unlike many other shell nouns discussed in Schmid (2000), *bottom line* has its roots in American English. It seems a likely development that it will become one of the frequently used shell nouns in a variety of Englishes. Similar developments have been observed by Aijmer (2013) with respect to pragmatic markers: "'new' pragmatic markers (or uses of pragmatic markers) travel quickly to other varieties", or by Mair (2009: 22), who demonstrates that *be like*, an innovation in American English, is now reported as an innovation in Australian and Canadian English. As *that's the bottom line is (that)* is also one type of innovative construction that originates from American English, it will be worth exploring whether this construction will spread to other regional variations of English.

---

14. What remains to be discussed is an analysis of the apo-koinou construction "that's the X is (that)" from the point of view of "chiasmus", which is defined as "a rhetorical device that juxtaposes structures with mirror-image syntax" (Los 2015: 177; see Yoshino 1984 for a comprehensive view of rhetorical devices including chiasmus in Old English). Indeed the structure of "that's the X is (that)" exhibits an almost symmetrical appearance (see Shibasaki 2015a for a pilot study on the issue).

Acknowledgements

References

Aijmer, Karin. 2013. *Understanding Pragmatic Markers*. Edinburgh: EUP.

Andersen, Gisle. 2002. Corpora and the double copula. In *From the COLT's mouth ….. and others': Language corpora studies in honour of Anna-Brita Stenström*, Leiv E. Breivik & Angela Hasselgren (eds), 43–58. Amsterdam: Rodopi.

Blockley, Mary. 2001. Subordinate clause without ἀπὸ κοινοῦ in Old English verse, chiefly in *Beowulf* and chiefly *nu* and *swa*. *Studia Neophilologica* 73: 4–10.

Bolinger, Dwight L. 1987. The remarkable double *is*. *English Today* 9: 39–40.
  doi:10.1017/S0266078400002728

Bybee, Joan. 2010. *Language, Usage and Cognition*. Cambridge: CUP. doi:10.1017/CBO9780511750526

Bybee, Joan. 2015. *Language Change*. Cambridge: CUP.

Coppock, Elizabeth, Brenier, Jason, Staum, Laura & Michaelis, Laura. 2006. "The thing is, is" is no mere disfluency. *Berkeley Linguistics Society* 32: 85–96.

Corminboeuf, Gilles. 2012. Des *apo koinou* aux constructions louches. In *Le verbe en verve: Réflexions sur la syntaxe et la sémantique verbales*, Marleen Van Peteghem, Peter Lauwers, Els Tobback, Annemie Demol & Laurence De Wilde (eds), 215–231. Gent: Academia Press.

Croft, William. 2001. *Radical Construction Grammar*. Oxford: OUP.
  doi:10.1093/acprof:oso/9780198299554.001.0001

Curzan, Anne. 2012. Revisiting the reduplicative copula with corpus-based evidence. In *The Oxford Handbook of the History of English*, Terttu Nevalainen & Elizabeth C. Traugott (eds), 211–221. Oxford: OUP.

Du Bois, John W. 2007. The stance triangle. In *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction* [Pragmatics & Beyond New Series 164], Robert Englebretson (ed.), 139–182. Amsterdam: John Benjamins. doi:10.1075/pbns.164.07du

Du Bois, John W. 2014. Towards a dialogic syntax. *Cognitive Linguistics* 25(3): 359–410.
  doi:10.1515/cog-2014-0024

Erman, Britt & Warren, Beatrice. 2000. The idiom principle and the open choice principle. *Text* 20(1): 29–62. doi:10.1515/text.1.2000.20.1.29

Flowerdew, John & Forest, Richard W. 2014. *Signalling Nouns in Academic English: A Corpus-based Discourse Approach*. Cambridge: CUP.

Günthner, Susanne. 2011. *N be that*-constructions in everyday German conversation: A reanalysis of 'die Sache ist/das Ding ist' ('the thing is')-clauses as *projector phrases*. In *Subordination in Conversation: A Cross-linguistic Perspective* [Studies in Language and Social Interaction 24], Ritva Laury & Ryoko Suzuki (eds), 11–36. Amsterdam: John Benjamins.
  doi:10.1075/slsi.24.03gun

Günthner, Susanne. 2014. The dynamics of *dass*-constructions in everyday German interactions – A dialogical perspective. In *Grammar and Dialogism: Sequential, Syntactic, and Prosodic Patterns between Emergence and Sedimentation*, Susanne Günthner, Wolfgang Imo & Jörg Bücker (eds), 179–206. Berlin: De Gruyter Mouton.

Hayashi, Makoto. 2004. Projection and grammar: Note on the "action-projecting use of the distal demonstrative *are* in Japanese." *Journal of Pragmatics* 36(8): 317–344. doi:10.1016/j.pragma.2004.05.006

Hoffmann, Sebastian. 2005. *Grammaticalization and English Complex Prepositions: A Corpus-based Study*. London: Routledge.

Hopper, Paul J. 2004. The openness of grammatical constructions. *Chicago Linguistic Society* 40(2):153–175.

Hopper, Paul J. 2007. Linguistics and micro-rhetoric: A twenty-first century encounter. *Journal of English Linguistics* 35(2): 236–252 doi:10.1177/0075424207305307

Hopper, Paul J. 2011. Emergent grammar and temporality in interactional linguistics. In *Constructions: Emerging and Emergent*, Peter Auer & Stefan Pfänder (eds), 22–44. Berlin: De Gruyter. doi:10.1515/9783110229080.22

Hopper, Paul J. & Thompson, Sandra A. 2008. Projectability and clause combining in interaction. In *Crosslinguistic Studies of Clause Combining* [Studies in Language 80], Ritva Laury (ed.), 99–123. Amsterdam: John Benjamins. doi:10.1075/tsl.80.06hop

Horlacher, Anne-Sylvie & Pekarek Doehler, Simona. 2014. 'Pivotage' in French talk-in-interaction: On the emergent nature of [clause-NP-clause] pivots. *Pragmatics* 24(3): 593–622. doi:10.1075/prag.24.3.07hor

Jespersen, Otto. 1949. *A Modern English Grammar on Historical Principles, Part III: Syntax, Second Volume*. London: George Allen & Unwin.

Kellner, Leon. 1892. *Historical Outlines of English Syntax*. New York NY: Gordon Press.

Kerr, Betsy. 2014. Left dislocation in French. In *Perspectives on Linguistic Structure and Context* [Pragmatics & Beyond New Series 244], Stacey Katz Bourns & Lindsy L. Myers (eds.), 223–240. Amsterdam: John Benjamins. doi:10.1075/pbns.244.11ker

Lambrecht, Knud. 1988. There was a farmer had a dog: Syntactic amalgams revisited. *Berkeley Linguistics Society* 14: 319–339.

Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39: 463–516. doi:10.1515/ling.2001.021

Longman Dictionary of Contemporary English, 5th edn. 2012. London: Longman.

Los, Bettelou. 2015. *A Historical Syntax of English*. Edinburgh: EUP.

Mair, Christian. 2009. Corpus linguistics meets sociolinguistics. In *Corpus Linguistics: Refinements and Reassessments*, Antoinette Renouf & Andrew Kehoe (eds), 7–32. Amsterdam: Rodopi. doi:10.1163/9789042025981_003

Massam, Diane. 1999. *Thing is* constructions: The thing is, is what's the right analysis? *English Language and Linguistics* 3(1): 335–352. doi:10.1017/S136067439900026X

Meritt, Herbert D. 1967. *The Construction ἀπὸ κοινοῦ in the Germanic Languages*. New York NY: AMS Press.

Miura, Ayumi. 2009. *Juliana* 329b-30a revisited: *Apo koinou* or not? *Studies in Medieval English Language and Literature* 24: 1–19.

Online etymology dictionary. <http://www.etymonline.com/>

Oxford Advanced Learner's Dictionary, 8th edn. 2010. Oxford: OUP.

Ross-Hagebaum, Sebastian. 2005. The *that's X is Y* construction as an information-structure amalgam. *Berkeley Linguistics Society* 30: 403–414.

Schmid, Hans-Jörg. 2000. *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Berlin: Mouton de Gruyter. doi:10.1515/9783110808704

Shapiro, Michael & Haley, Michael C. 2002. The reduplicative copula *IS IS*. *American Speech* 77(3): 305–312. doi:10.1215/00031283-77-3-305

Shibasaki, Reijirou. 2014a. On the development of *the point is* and related issues in the history of American English. *English Linguistics* 31(1): 79–113.

Shibasaki, Reijirou. 2014b. On the grammaticalization of *the thing is* and related issues in the history of American English. In *Studies in the History of English Language VI*, Michael Adams, Robert D. Fulk & Laurel J. Brinton (eds), 99–121. Berlin: De Gruyter Mouton.

Shibasaki, Reijirou. 2015a. Chiasmic constructions in Present-day English. Ms, Meiji University.

Shibasaki, Reijirou. 2015b. Gendai amerika eigo no nijuu kopyura koobun (Double copula constructions in Present-day American English). In *Nichi-Eigo no Bunpooka to Koobunka* (Grammaticalization and constructionalization in Japanese and English), Minoji Akimoto, Hirofumi Aoki & Mitsuru Maeda (eds), 147–180. Tokyo: Hituzi Syobo.

Traugott, Elizabeth C. 2014. Toward a constructional framework for research on language change. In *Grammaticalization – Theory and Data* [Studies in Language Companion Series 162], Sylvie Hancil & Ekkehard Köning (eds), 87–105. Amsterdam: John Benjamins.

Traugott, Elizabeth C. & Trousdale, Graeme. 2013. *Constructionalization and Constructional Changes*. Oxford: OUP. doi:10.1093/acprof:oso/9780199679898.001.0001

Tuggy, David. 1996. The thing is is that people talk that way: The question is Why? In *Cognitive Linguistics in the Redwoods*, Eugene H. Casad (ed.), 713–752. Berlin: Mouton de Gruyter. doi:10.1515/9783110811421.713

Ukaji, Masatomo. 2003. Subject zero relatives in Early Modern English. In *Current Issues in English Linguistics*, Masatomo Ukaji, Masayuki Ikeuchi & Yoshiki Nishimura (eds), 248–277. Tokyo: Kaitakusha.

Visser, Fredericus. Theodorus. 1963. *An Historical Syntax of the English Language, Part one: Syntactical Units with One Verb*. Leiden: E.J. Brill.

Wray, Alison. 2009. Identifying formulaic language: Persistent challenges and new opportunities. In *Formulaic Language, Vol. 1: Distribution and Historical Change* [Typological Studies in Language 82], Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali & Kathleen M. Wheatley (eds), 27–52. Amsterdam: John Benjamins. doi:10.1075/tsl.82.02ide

Yaguchi, Michiko. 2010. The historical development of the phrase *there's*: An analysis of the *Oxford English Dictionary* data. *English Studies* 91(2): 203–224. doi:10.1080/00138380903355007

Yoshino, Yoshihiro. 1984. Poetic Syntax in the Old English "Meters of Boethius": A Comparative Study of Verse and Prose. PhD dissertation, University of Pennsylvania.

Corpora

The British National Corpus (*BYU-BNC*), 100 mil. words. <http://corpus.byu.edu/bnc/>

The Corpus of Contemporary American English 1990–2012 (*COCA*), 450 mil. words. <http://corpus.byu.edu/coca/>

The Corpus of Historical American English (COHA), 1810–2000s, 400 mil. words. (http://corpus.byu.edu/coha/)

The Oxford English (OED), second edition on CD-ROM Version 4.0, Oxford: OUP.

The Santa Barbara Corpus of Spoken American English (*SBCSAE*), c. 249,000 words, Department of Linguistics, University of California, Santa Barbara. <http://www.linguistics.ucsb.edu/research/santa-barbara-corpus>

Index

The papers in this volume cover a wide range of interrelated syntactic phenomena, from the history of core arguments, to complements and non-finite clauses, elements in the clause periphery, as well as elements with potential scope over complete sentences and even larger discourse chunks. In one way or another, however, they all testify to an increasing awareness that even some of the most central phenomena of syntax – and the way they develop over time – are best understood by taking into account their communicative functions and the way they are processed and represented by speakers' cognitive apparatus. In doing so, they show that historical syntax, and historical linguistics in general, is witnessing a convergence between formerly distinct linguistic frameworks and traditions. With this fusion of traditions, the trend is undeniably towards a richer and more broadly informed understanding of syntactic change and the history of English. This volume will be of great interest to scholars of (English) historical syntax and historical linguistics within the cognitive-linguistic as well as the generative tradition.

JOHN BENJAMINS PUBLISHING COMPANY