# Formal and quantitative approaches to the study of syntactic change: three case studies from the history of English

Zimmermann, Richard



This publication URL: https://archive-ouverte.unige.ch/unige:96500
Publication DOI: 10.13097/archive-ouverte/unige:96500

UNIVERSITÉ DE GENÈVE
FACULTÉ DES LETTRES

# Formal and Quantitative Approaches to the Study of Syntactic Change: Three Case Studies from the History of English

THÈSE DE DOCTORAT

présentée à la

Faculté des lettres
de l'Université de Genève

pour l'obtention du grade de
Docteur ès lettres

par

RICHARD ZIMMERMANN

April 25, 2017

Directeur de thèse

ERIC HAEBERLI
Professeur Associé, Département de Langue et Littérature Anglaises, Université de Genève

Présidente du jury

GENOVEVA PUSKAS
Professeur Associé, Département de Langue et Littérature Anglaises, Université de Genève

Membres du comité de défense de thèse

ANTHONY KROCH
Professor of Linguistics, University of Pennsylvania

SUSAN PINTZUK
Professor, Department of Language and Linguistic Science, University of York

KERSTI BÖRJARS
Professor, School of Arts, Languages and Cultures, Department of Linguistics and English Language, The University of Manchester

To my friends.

# Acknowledgements

*nani gigantum humeris insidentes*
*'dwarves standing on the shoulders of giants'*

– Bernard of Chartres, c. 1120

Writing this PhD dissertation has been one of the most tremendous experiences of my life. Oscillating between episodes of euphoria and fits of frustration, I was always accompanied by extraordinary people who supported, encouraged and advised me. I am truly grateful to all of them from the bottom of my heart!

First and foremost, this work would have been entirely impossible without the help, guidance and trust of my thesis adviser, Eric Haeberli. I have enjoyed exceptional freedoms during my time as a PhD student. No topic was too off-limits, no idea too absurd, for intensive exploration and discussion. I could use LFG for my formalizations even though this framework is not standardly taught at our institution. I was even allowed to veer into a completely new language for some time and construct a small corpus of Old High German. Thank you so much for absolutely everything!

Number two on my list of people vital for this project is my immediate family. Nothing gives me more strength in my life than to know that I can always count on my dad, Olaf, my mom, Birgit, and my sister, Anne. Hab' euch total lieb!

My work profited greatly from the input provided by numerous scholars, professors, PhD students, conferences audiences and colleagues. While I cannot remember each and everyone of them, I would like to thank at least some people explicitly.

I have enjoyed an excellent work atmosphere with my office mates, Margherita Pallottino, Dara Jokilehto, Ji Young Shim and Tabea Ihsane.

Thanks are due to Paola Merlo, Genoveva Puskas, Ur Schlonsky, and many other linguists at the University of Geneva who have given me feedback or discussed linguistic issues with me,

to our secretary, Eva Capitao, for helping me out of hairy situations countless times and for reading the horoscope with me every day to practice some French,

to Tony Kroch, Charles Yang, Beatrice Santorini, Don Ringe and other scholars at the University of Pennsylvania, who spent a lot of time teaching me, discussing my research and providing me with invaluable feedback during my academic year abroad at UPenn in 2014/2015,

to Ans van Kemenade and Meta Links for making me feel very welcome during a productive, two-week stay at Radbound University in August 2015,

to Susan Pintzuk, Ann Taylor, Anthony Warner, Peter Sells and other researchers at the University of York for their encouragement and for accepting me as a visitor during a month-long stay in July 2016,

to Cynthia Allen for sending me a lot of useful material on Old English syntax and text dating,

to George Walkden, Kersti Böjars, Robert Truswell, Hélène Bellon Méguelle, Anamaria Bentea, Léna Baunaz, Kristina Gulordava, Yves Scherrer, Sharid Loáiciga, Giuseppe Samo, Joel Wallenberg, Peter Petré, Manuela Schönenberger, Frederick Newmeyer, Martin Hilpert, Nigel Vincent, Sali Tagliamonte, and everybody else who has engaged with me in constructive, critical or challenging conversation.

*hu sliþen bið*
*sorg to geferan*
*þam þe him lyt hafað*
*leofra geholena*

*'how cruel is*
*sorrow as a companion*
*to the one who has few*
*beloved friends'*

– *The Wanderer* (lines 29-31, from a late 10th century manuscript)

Yet, the people who had the most profound influence on me during my time as a PhD student were not my colleagues or my family, but my friends. They kept me sane. They made my life rich, enjoyable and exciting. They taught me that there are more important things than to spend the weekend in the office trying to read an obscure Anglo-Saxon saint's life (though maybe not much more important).

I want to give a shout-out to my befftest friend in the whole world, Beff, just for being awesome, Neele, and their son, Taro, who, although still too young to know if we're friends, I hope will become one in the future.

Steffen Eisner is a perpetual source for encouragement and even read substantial parts of this thesis providing me with very useful comments.

Then there is the 'Geneva Crew' - Matthew Brouillard, Christian Morreira, Jacob Strand, Julien Levy and Nicolas Hinz. Thanks so much for all the fun times, guys! In Mareike Lagoda, Eva Soltész and Christoph Spennemann I found great friends in our university choir. Several people have passed through Geneva, and in doing so made my life richer, especially Douglas Armendone, Nate King, Sébastien Peytrigny, Patrice Good and Joel Anderson. Tim Caron turned from my tandem partner into a valued friend. The best house parties in Geneva are clearly thrown by Laura Schmed and Marc Baumgartner. Rachel Nye became a dear friend during her research stay in Geneva in 2012.

I count among some of my closest friends people I met during my time as an undergraduate student in Heidelberg. Most importantly, I'm proud to be friends with Moritz and Lindsey Schuler. They are full of life and an inspiration! I wish I had had a bit more time to spend with my Heidelberg fraternity brothers during those last years, but I'm sure I can catch up in the future, especially with Olaf Giebel, Rochus Hermann, Martin Höcker, Steven Punt, their families, but also everybody else. I was lucky to have become friends with a particularly nerdy group of linguistics students and lecturers during my time in Heidelberg - Mike Schiffmann, who has an incredible knowledge about the history of Generative Grammar, Ivo Ivanov, Neven Wenger and Benjamin Börschinger, who has given me an incredible amount of food for thought over the years.

I'm very grateful that I could make some amazing friends in the incredibly vibrant postgraduate linguistics community at Penn, especially Milena Šereikaitė, Robert Wilder, Betsy Sneller, Aaron Ecay, Rodrigo Gier, Duna Gylfadóttir, Ava Irani, Brittany Dael McLaughlin, Sabriya Fisher, Aaron Freeman and Luke Adamson. Many of them were kind enough to provide me with grammaticality judgments. I also met wonderful people outside of the classroom, particularly Matthew Steele, Joe Sengoba, Andrea Smith, Marcos Gonzalez Palma, Jeff Schwane, John Ford and Joel Fineman.

There are many more people whose friendship I would not want to have missed during those last years, Bruno Richard, Russ Norton, Michellie Young, Buschman, my cousins Anita, Felix, Franziska, Hans, Judit, Martin, Max Zimmermann, Yves Roulet, João Ribeiro, Marc Benz, Patrice Good, Quentin Markarian, James Welch, Pete Cummings, and whoever else I may, accidentally, have missed.

I'm unspeakably happy that you're all in my life. And so, dear friends, liebe Freunde, cher(e)s ami(e)s – this thesis is for you.

# Contents

# List of Figures

# List of Tables

# 1

# Studying Syntactic Change - Preliminaries

*"To study history means*
*submitting to chaos*
*and nevertheless retaining faith in order and meaning."*
– Hermann Hesse *The Glass Bead Game*, 1943

## 1.1 Introduction

The essence of language has been described in terms of many metaphors. Saussure compared the gradual unfolding of linguistic rules, their conventionalized application and interconnected character to a game of chess (Saussure 1916/2014: e.g., part 1, chapter 3, §4). Baker likened the limited scope of linguistic entities to combine into incredibly diverse syntactic structures to the small number of chemical elements in the periodic table and their potential to form vastly different molecules (Baker 2001). Bloomfield equated the difference between a partial, conscious knowledge of linguistic structures and the complexities of truly mastering to speak natively with the difference between learning about musical theory, notes, intervals or chords, and actually creatively performing a musical production (Bloomfield 1942).

I have always preferred a conceptualization of language in terms of living organisms. Just like species are a collection of individual animals, languages are produced by individual speakers. Biologists usually define members of a species as individuals that can produce fertile offspring; linguists commonly regard different dialects as instantiations of the same language if they are mutually intelligible, and both fields struggle with their definitions in the form of ring species or dialect continua. An organism's genotype and phenotypic expression can be compared to a grammar's set of morphemes and combinatory rules, or competence, and the actual usage of this underlying potential, or performance. Animals may leave fossils behind as imperfect and fragmentary expressions of their genome; speakers can produce text witnesses as imperfect and fragmentary reflections of their mental grammar. But most importantly, just as "nothing in biology makes sense except in the light of evolution" (Dobzhansky 1973), so, too, nothing about the form and function of linguistic structures truly make sense except in terms of their history. With an outlook like this, it should not come as a surprise that my foremost linguistic interest is in language change, the gradual transition from one language stage into another.

The purpose of this thesis is (i) to explore general mechanisms of language change in general and (ii) to contribute to a better understanding of certain facets of the history of the English language in particular. I will present three cases studies on the development of syntactic structures in historical English. Each of them will investigate the diachrony of one particular syntactic phenomenon, formalize it in a grammar model that incorporates variable and competing grammatical options, derive from that model quantitatively testable hypotheses, and subsequently examine these predictions in a series of corpus studies. In this way, I hope to illustrate how formal and quantitative approaches to syntactic change can be united in a productive and elegant manner.

## 1.2 Central Notions in the Study of Syntactic Change

I will examine a few critical concepts for the study of syntactic change. First, I will present a simple view on the changeable elements of a mental grammar in terms of lexical items and syntactic rules. Subsequently, I will introduce the concepts of competition and common replacement changes. Finally, I will discuss the relevance of s-shaped curves and the Constant Rate Hypothesis for the study of language change.

### 1.2.1 Basic Linguistic Units

In order to grasp, at least in highly idealized terms, the fact of language change, one must probably assume that there exist, for the purpose of inheritance, some primitive, i.e., not further decomposable, linguistic units that can be passed on from one generation to the next. They are, in metaphorical terms, the genes of a language. Such linguistic replicators have been given many different labels, notably, adjustable 'parameters' (e.g., Lightfoot 1999: 59), a special notion of 'grammar' (e.g., Kroch 1989, Roeper 1999, Yang 2002), 'abstract grammatical option' (Kroch 1994), 'units and principles [that] may be identified as a linguistic feature' (Mufwene 2001: 2), 'constructions' contained in a network called the 'construct-i-con' (Goldberg 2003), the 'lingueme' (Croft 2000), and so on and so forth.

There are two different types of grammatical elements in the ontology of linguistic units of inheritance used for the case studies of this dissertation. They will be formalized with the tools and technologies made available in the framework of Lexical-Functional Grammar (LFG) (Bresnan 2001).

On the one hand, I will make use of the traditional concept of a 'lexical item,' defined as the triplet of a signifier, a category and a bundle of attribute-value paired features that arise through structuralist differentiation (e.g., English 'singular' gains meaning by the fact that it is not 'plural'). I assume that lexical items function as the basic input to syntactic computation (following some form of lexicalism). An example of a lexical item formalized in LFG is shown in (1). The word has the form *houses*, is of category `N` and will distribute like other common, count, plural nouns.

(1) *houses* `N` (↑ PRED)= 'HOUSE'
(↑ N-TYPE)= common-count
(↑ NUM)= pl
(↑ PERS)= 3

On the other hand, I will employ context-free phrase structure rules annotated with functional constraints that regulate grammatical dependencies. An example of such a phrase structure rule is shown in (2). It asserts that a determiner phrase, `DP`, can be re-written to a head, `D`, and a noun phrase, `NP`. Further, the features of the head's attribute-value matrix, denoted by its metavariable ↓, and the features of the `NP`'s attribute-value matrix, denoted by its metavariable ↓, are unified, as shown by the fact that both point to same mother node, `DP`, denoted by the metavariable, ↑. This rule can then generate, i.e., algorithmically enumerate, all nominal constituents of the corresponding structure in a language, like *the houses*, *my children*, etc.

(2) $\texttt{DP} \rightarrow \begin{array}{c}\texttt{D}\\ \uparrow = \downarrow\end{array} \quad \begin{array}{c}\texttt{NP}\\ \uparrow = \downarrow\end{array}$

The totality of lexical items and phrase structure rules form a mental grammar. My working assumption is that these grammar components are the primitives of language acquisition, by extension, of linguistic inheritance and change, and that these units are acquired in their entirety as discrete units. There are many problems that immediately arise from such a conceptualization of grammar. The most pressing questions would probably be (i) how to make the grammar components sensitive to register and socio-linguistic contexts, (ii) how to represent their probability of use or entrenchment in the mind etc., (iii) how to lexicalize their applicability in order to model collocates, idioms or multi-word constructions, and (iv) how to incorporate morphological processes of word formation and inflection as inheritable rules. I believe that these questions can be left open here and do not obstruct fruitful analyses of the specific language changes in my case studies. My definition of grammar is thus a rather great abstraction and idealization of reality, but one that is appropriate for the purpose of this work.

### 1.2.2 Grammar Competition and Common Replacement Changes

It is possible to distinguish between two different kinds of linguistic changes based on the way they are propagated in a population. (i) The first kind could be called an 'intrusive change.' It is characterized by a new form entering (or leaving) the language system and starting to spread (or decline). Such a change adds to (or subtracts from) the expressive repertoire of a language. Its rise can be measured in texts as its normalized occurrence per a certain number of words. One could compare it to the spread of some fortuitously advantageous mutation (or the loss of a gene) in biology. For example, the expression *end up doing something* did not exist before c. 1930 but has steadily gained ground since then and has now reached quite a substantial level of use (based on a Google books n-gram search). (ii) The second kind of change can be labeled a 'competitive change.' It proceeds in the form of variants of a linguistic variable such that, over time, an innovative variant drives out a conservative alternative. This type does not change the expressive power of a language but merely replaces one grammatical option with a (near-)synonymous one. Its development can be measured as the proportion of the innovative variant out of all variants. A biological analogy is afforded by the existence of alleles, i.e., alternative forms of a gene. For instance, Middle English first and second person possessive determiners contained an *n*-sound, *mine*, *thine*, but were ousted by *n*-less allomorphs, *my*, *thy*, during Early Modern English (Hilpert 2013: ch. 3).

Competitive changes have often been described in terms of so-called 'grammar competition' (e.g., Kroch 1989). The name references two states of a grammar, an earlier one that does not include the innovative variant, and a later one that does. Competition leads to a transitional period between the two states. Grammar competition has invited criticism from scholars who interpret the concept to imply a mixture of entire dialects or a juxtaposition of holistic mental systems, and condemn such a notion as an "unrealistic model of variation" (Denison 2001: 120), as a "sort of 'last resort' hypothesis, which can only be invoked when we have already tried to make sense of the data on the basis of a single grammar" (Poletto 2014: 172) or as inferior to the assumption "that grammar competition involves competing parametric options within a single - instead of multiple - grammars" (Fuss 2017: 477, fn. 27), etc. It is immaterial here whether this kind of criticism adequately reflects or misrepresents the original intent of the concept and whether it is justified or not. I will simply appropriate the label 'grammar competition' for my own purposes. I will use the term 'grammar competition,' and related expressions, such as 'syntactic rules compete' or 'competing forms,' etc. in the sense of a 'competitive change' as outlined above, i.e., as competing variants of a dependent, linguistic variable.

Within the group of changes that can fruitfully be analyzed as grammar competition, one can, I think, make further valuable distinctions. One could differentiate between changes involving just two variants, one conservative, one innovative, and those with more than two variants. It would also be possible to distinguish between complete changes, in which an innovative variant completely drives out its competitor(s), and failed changes, in which an innovation starts to take hold but is later reverted and abandoned. Further, there may be changes in which the variants undergo conditioning or specialization, so that each will eventually be used only in their own, complementarily distributed contexts. Complex changes of these kinds seem to me to be relatively ill-understood at the moment. It seemed better to me to investigate and learn from elementary cases of syntactic change first and leave work on the above-mentioned, more complex kinds of competitive changes for the future. I therefore decided to limit the empirical material of my case studies to instances of linguistic changes that are as straightforward as possible. For this purpose, I would like to introduce the concept of a 'common replacement change.'

Common replacement changes are competitive changes that display the three defining characteristics binarity, completeness and measurability.

- Common replacement changes are binary. They proceed via competition between exactly two variants, one conservative, one innovative. The variants should be easily identifiable by their form.
- Common replacement changes are complete. They involve perfect substitution of the conservative by the innovative variant. Hence, the innovative form does not trail off at some level of use; the change does not fail; the changing forms do not become conditioned on a contextual factor, etc. Ideally, the increasing proportion of the innovative form can be observed from 0–100% of use in the textual record, but the initial or final stages may not be attested.
- Common replacement changes are measurable. There is a textual corpus that allows tracking the implementation of the change adequately through history. There should be no evidence of prescriptive pressure, censorship or other cultural biases against the innovative form. The text sample should not be irreparably harmed by scribal interference, copying errors, distortive dialect and genre mixing, non-native speaker intrusion, forgeries or other factors undermining representativeness. The text record should be large. This is, of course, a subjective assessment, but generally speaking a change taking effect over a few centuries should be quantified with hundreds, better thousands, of examples, so that isolated examples are unlikely to substantially alter the overall conclusions.

I chose the attribute 'common' for the term because common replacement changes seem to occur quite frequently, because they have often been analyzed before, e.g., in the variationist paradigm of socio-linguistics, and because they are relatively simple by comparison with non-binary or incomplete changes.

The definition of a common replacement change now allows a simple model for the implementation of a linguistic innovation in a group of speakers. One can simply say that (i) grammars of the conservative period only represent the conservative variant, (ii) the actuation of a change is the addition of a competing, innovative variant into a mental grammar, (iii) both the conservative and the innovative variant can produce linguistic output during a transitional period, in which the innovative period steadily gains ground, and (iv) eventually grammars of the innovative period can only make use of the innovative variant whereas the conservative variant has become unproductive. This is illustrated below.

(3) *Implementation of a common replacement change in a population*

| | | |
|---|---|---|
| Conservative period: Speakers represent only the conservative variant $\alpha$ | Transitional period: Speakers represent both the conservative variant $\alpha$ and the innovative variant $\beta$ | Innovative period: Speakers represent only the innovative variant $\beta$ |

$$\alpha \quad \rightarrow \quad \{ \quad \alpha \quad | \quad \beta \quad \} \quad \rightarrow \quad \beta$$

### 1.2.3 S-Shaped Curves and the Constant Rate Hypothesis

Grammar competition is usually discussed in connection with two related aspects: the rise of linguistic innovations in the form of s-shaped curves and the Constant Rate Hypothesis. The two concepts will be of relatively great importance in this dissertation.

The diffusion of a linguistic form is usually described as a saturation process. Hence, the spread of a new linguistic form should occur slowly at first, then speed up as more and more speakers begin to adopt the innovation and use it more frequently, and eventually slow down again as the innovative variant reaches full penetration in the population. In other words, the implementation of a linguistic innovation should obey logistic growth assumptions and thus follow an s-shaped curve, as observed, say, for the dissemination of a new technology or product in a market, or for the development of bacterial cultures in a petri dish, etc., Logistic, s-shaped growth is the most common model of the implementation of a linguistic change (e.g., Weinreich et al. 1968), but some scholars have warned against over-generalizing its application (Denison 2003).

One common way to model the assumed s-shaped trajectories of linguistic innovations makes use of the logistic function. It characterizes a probability distribution, i.e., it ranges from 0 to 1, and has a sigmoid shape. Fitting a logistic function through a data cloud of observed innovative and conservative examples from corpora can be achieved with logistic regression. Most importantly, the fitted function will include a coefficient for the time variable, which determines the angle or slope of the increase and hence quantifies the speed with which an innovative variant spreads through a population. This parameter can therefore be called the 'rate of change.' The statistical techniques required for logistic regression are standardly available in modern statistics packages. I will use the free statistics software *R* (Bates et al. 2008) to build statistical models of the changes examined in my case studies.

It is often instructive to split up the data collected for a linguistic change into linguistically meaningful subgroups. These subgroups are referred to as 'contexts.' Contexts are, in principle, limited only by the creativity and interests of the investigator. For instance, the appearance of an innovative form could be hypothesized to be influenced by the the kind of phoneme that follows it (phonological context), by the clause type in which it is positioned (syntactic context), by the gender or social class of the speaker (socio-linguistic context), by the nearby presence or absence of a semantic feature, say, animacy (semantic context), by the distance to the last mention of the item or a related item (pragmatic context), and so on.

However, I am interested in one special kind of context. For the determination of these particular contexts, one first has to have a fairly clear model or theory of the grammatical phenomena under scrutiny. Specifically, one can allow a number of preexisting rules in the set of production rules of a grammar to utilize an innovative or conservative grammatical element under investigation. Application of different rules to the innovative or conservative variant will thus trivially result in distinct output. These outputs, then, characterize the relevant contexts for this dissertation - they are distinct structural representations that are attributable to distinct applications of grammatical rules relative to a grammar model or theory. They can be regarded as special in that they are “tied together grammatically” (Kroch 1989: 24) or as being determined by a “shared syntactic structure” (ibid.: 201) (see Corley (2014) for criticism regarding the concept of grammatical contexts).

Now, the innovative variant may be more or less likely to apply to some rules than to others. This means that as one measures the change in corpora and subdivides the retrieved examples into different contexts of the relevant kind, one may find that certain contexts prefer, others disprefer, the occurrence of the spreading innovative form. However, it has been hypothesized and subsequently repeatedly observed that the rate of change, the speed at which the innovative variant establishes itself, will still be identical in all grammatical contexts. The prediction of identical rates of change between different grammatical contexts is known as the Constant Rate Hypothesis (Kroch 1989). Logically, the Constant Rate Hypothesis can only be operative if the rules that access the innovative variant are not themselves in the process of changing, but remain stable. In other words, the effects of the linguistic contexts are independent of time. Practically speaking, the Constant Rate Hypothesis claims that characteristics of the contexts themselves, say their frequencies, their complexity, their perceived prestige, etc., will be wholly irrelevant for the rate at which a new form gets implemented in a population (see for example Kroch’s (1989) discussion especially with reference to an alternative model of change proposed by Charles Bailey). The formulation in (4) presents my own understanding of the Constant Rate Hypothesis in terms of a conditional statement.

(4) *The Constant Rate Hypothesis*

If:
- an innovative element enters a language and gradually replaces a conservative competitor,
- and a number of different grammatical rules can be applied to the innovative element creating different contexts
- and each of these grammatical rules remains stable through time,
- and data is collected for the different contexts from a representative corpus in numbers sufficiently large for precise statistical estimates,

Then:
→ one will observe that the rate of change for the spread of the innovative element will be identical in all contexts.

The Constant Rate Hypothesis is of great significance for the identification of the correct abstraction level of a change. If different contexts change at the same rate, this suggests that they involve the same underlying changing lexical item or grammar rule. The discovery of a constant rate effect in a data set thus allows to point very precisely to the changing unit.

## 1.3 Why do Languages Change?

The question of *why* a linguistic phenomenon would change to begin with, rather than stay stable, at the time that it does, and not later or sooner, is particularly prominent in historical linguistics. Unfortunately, the field still seems to be far away from providing a comprehensive, evidence-based and principled answer to this long-standing question. I would like to mention two basic problems in establishing causal associations for language change.

Firstly, it is not really clear what is meant by ‘why.’ What kind of answer would be expected? What would be a criterion for the goodness of an offered explanation? “Languages change because of language-external developments, invasions, migrations, technological advances” - is this a satisfactory answer? “Languages change because the featural make-up of lexical items can vary” - does this statement have explanatory value? At what point should linguists be satisfied with a proposed causal link between one variable and a linguistic change?

Secondly, explanations often seem to be advocated from a particular, relatively subjective vantage point. This seems to be a characteristic partially shared with history as an academic discipline. For instance, when one looks for answers to the question “Why did the Roman Empire collapse?”, one will find that every answer is influenced by individual convictions and concerns by the spirit of the age. In the early $20^{\text{th}}$ century, after the beginning repeals of anti-miscegenation laws, the fall of Rome was seen as a consequence of race mixing (Tenney 1916). In the 1980s, during a time when the Iran-Contra affair revealed massive corruption in the Reagan administration, Rome’s collapse was attributed to the corrupt, morally bankrupt nature of her officials (MacMullen 1988). In the 2000s, at a time of multiculturalism, globalization and refugee crises, the decline of the Roman Empire was connected to mass migration and a loss of cultural identity (Ward-Perkins 2006). Similarly, answers to the *why*-question of language change can be expected to be fundamentally different if they are given by a usage-based functionalist, who wants to eliminate the need for a language-specific mental grammar, or by a generative formalist, who wants to attribute variation to minimally different parameters on an otherwise invariant Universal Grammar, or by a traditional philologist, who wants to connect texts to their manuscript witnesses and external history.

Both of these problems can be mitigated, even though not completely overcome, by employing a sound, scientific methodology. For me, that implies a two step process of guessing and experimentation. For the first step, all conceivable explanations can be put on the table, no matter what kind of ontology and theory of causation they presuppose, the first problem, or what kind of subjective biases may have led to their proposition, the second problem. The process simply involves the identification and elucidation of a causal factor from an infinite space of possible explanations. The second step consists of deriving and testing hypotheses that should be true if the proposed cause is true but that will be false otherwise. Such hypothesis tests can then be used to decide between competing theories. It is a nontrivial, challenging, creative task to establish relevant conjectures that have the potential to falsify a hypothesized causal link. A theory can then be said to be the better affirmed the more predictions it makes and, conversely, the more falsification attempts it has survived. It seems appropriate to me to repeat these precepts of good science in connection with the *why*-question of change because my view is that they are often ignored or followed quite superficially and haphazardly while they are simultaneously absolutely essential if progress towards answering the question is to be made.

Finally, I will present a short outline of different kinds of explanations for linguistic change that have been presented in the specialist literature.

(1) First, there are approaches that use evolutionary dynamics to account for language change. Such work has, in my opinion, the greatest potential for a substantive understanding of general mechanisms of change. Essentially, evolutionary approaches try to identify a linguistic analogue to natural selection in biological evolution. This entails the identification of an advantage that new, intruding forms have over their conservative competitors. Mathematical replicator-mutator equations can subsequently be used to model stability, random drift and s-shaped, deterministic selection. The question then becomes how the advantage of a form should be determined and measured. A host of different suggestions have been presented.
(a) There are several abstract, formal definitions of linguistic fitness. Yang (2002) offers a definition of a grammatical option's advantage in terms of their relative copying fidelity, or recognizability. Jäger (2004, 2008) presents a game theoretic system of differential payoffs for different linguistic strategies optimizing communication between speaker and hearer.
(b) Innovative forms have also been analyzed as superior to their competitors on account of reduced structural complexity (e.g., Clark and Roberts 1993). In the Minimalist Program (Chomsky 1995), for instance, such considerations are frequently formulated in terms of comparisons between alternative derivations and preferences or economy conditions imposed on them (see e.g., the papers collected in Wilder (1997*b*)).
(c) Some researchers focus on functional reasons that may make one form more suitable than another in a given context. They study, for instance, effects of frequency of use contributing to phonological attrition and grammaticalization (e.g., Peng 2012), discourse and pragmatic factors in aiding processing, production and parsing (e.g., Wasow and Arnold 2003), memory constraints influencing word order (e.g., Hawkins 1994), and generally links between usage and cognitive storage, sometimes to the point of denying the existence of a mental grammar altogether (e.g., Bybee 2012).
(d) Finally, there may be cases in which one variant is consciously preferred over another one. In fact, some scholars ascribe to socio-linguistic preferences the primary advantage of an incoming form over a conservative one: "[T]he selection process is essentially a social one [...]. The variants in a linguistic variable have social values associated with them. Speakers select variants to use - that is, to replicate - in particular utterances on the basis of their social values: overt or covert prestige, the social relation of the speaker to the interlocutor, etc." (Croft 2000: 32).
It is reasonable to assume that all of the above factors (and possibly others) can contribute in varying degrees to the advantage of a linguistic variant. The multitude of possible fitness criteria may be one of the reasons why an adequate theory of language change in an evolutionary paradigm has not yet been formulated.

(2) In contrast to considerations of replication and selection, some accounts of linguistic change content themselves with listing observable diachronic patterns in an attempt to present a relatively theory-neutral taxonomy of linguistic changes. For instance, the Neogrammarian dictum of regular sound change and analogical leveling is a powerful, standardized mechanism of change and has led to several extensions, such as universal laws of analogy (for an excellent summary, see Trask 1996: 105-115). Linguistic change can also be classified according to general principles of internal, social or cognitive factors (Labov 1994, 2001, 2010).

(3) Finally, it seems relevant to at least list the causal influence from the external world on linguistic systems. As societies change, so, too, do languages.

## 1.4 Outline of the Thesis

This thesis consists of three case studies that try to shed light on research questions related to the issues of language change discussed in the preceding sections.

Chapter 2 will report one of the most careful and comprehensive tests of the Constant Rate Hypothesis to date. It is designed to correspond closely to my theoretical understanding of the concept. That is to say, I will identify a common replacement change, ensure that the antecedent conditions of my formulations of the Constant Rate Hypothesis in (4) are met, especially with respect to the deduction of contexts from a formal grammar model and the measurability of the change in a large corpus, and finally test if the consequent clause of the formulation also holds, i.e., if I can observe identical rates of change in each context. The empirical material for the chapter concerns a change in the realization of possessive *have* in late Modern American English, from c. 1810–2009. The purpose of the chapter is to put the basic assumptions of syntactic change from section 1.2 on a sound footing.

Chapter 3 subsequently deals with a much more complex research question - the detection of interactions between syntactic changes. The study presupposes the correctness of the Constant Rate Hypothesis. However, it investigates a case in which one of the antecedent conditions for its applicability is false, namely the requirement that the grammatical contexts themselves remain stable and thus be independent of time. The empirical material for the chapter consists of changes in verb placement and topicalization in Early English, from c. 800-1200. They are measured in a diachronically unstable context, a special clause type that is characterized by introductory conjunctions. One would thus predict violations of constant rate effects with respect to the changes in ordinary main clauses vs. main clauses that are introduced by a conjunction. The chapter attempts to make a meaningful contribution in that it presents a novel way to test the Constant Rate Hypothesis and simultaneously assigns a new use to it.

Chapter 4, finally, is specifically dedicated to an exploration of the *why*-question of change. While I briefly consider causes for some linguistic changes in the previous chapters as well, only the final study will systematically and thoroughly apply a common replacement change to the general methodology for the investigation of causal influences outlined in section 1.3. The empirical material is afforded by the replacement of the adverbial clause subordinator *then* with *when* during the course of Middle English, c. 1000-1500. The data for the investigation is taken in part from a newly constructed parsed corpus of Middle English poetry. The merit of the chapter lies in problematizing some central issues in the identification of causes for linguistic changes by reference to a practical example.

A short conclusion will round off my discussions.

# 2

# A Simple Syntactic Change in Modern Times: The Realization of Possessive *have* in American English[1]

*"We must learn to walk before we can run"*

– *Traditional English proverb*

## 2.1 Introduction

The purpose of this chapter is to present a new test case for the Constant Rate Hypothesis (CRH) (Kroch 1989), the idea that a linguistic innovation spreads at identical rates in all environments in which it is used, and to thereby substantiate the assumption of its correctness for subsequent chapters. The study will be different from previous investigations of the CRH in two important respects. Firstly, the relevant environments are not arbitrary divisions but deduced from a standard grammatical model of Modern English clause structure. It is thus largely uncontroversial that the variations in the different contexts do in fact reflect the same underlying linguistic change. Secondly, the data set I collected is very large, which allows for unprecedented statistical precision. These two advantages should render the outcome of this study relatively persuasive and deflect some of the concerns about the applicability of the CRH that have been raised in the past (e.g., Corley 2014, Denison 2003, Janda and Joseph 2003*a*: fn. 21).

The empirical material for this chapter concerns possessive *have* in late Modern to Present-Day American English. For the purpose of this chapter, I define "possessive *have*" by its typical complementation pattern. Instances of "possessive *have*" are all and only those structures that include a finite form of the verb *have* and an object DP denoting in a wide sense a possessed element. As such, the structure contrasts with other complementation patterns of *have*, such as the perfect auxiliary, which occurs with a past participle (*have done something*), causative *have*, which is followed by a non-finite clause with an overt subject (*have someone do something*), or its passive voice variant (*have something done*), modal *have*, which requires a *to*-infinitival clause (*have to do something*) etc. I do not impose any semantic requirements on the possessed DP (for a summary of semantic definitions, see Myler 2014). Hence, I indiscriminately consider all relations expressed by *have* between a possessor and a possessed element as valid instances of possessive *have*. This includes, but is not limited to,

[1]Parts of this chapter were presented on various occasions. I showed preliminary findings at the *VariaForMea* doctoral Winter School at Macolin, Switzerland, in December 2013. I am grateful for fruitful discussions with the other participants and the invited lecturer, Frederick Newmeyer, about whether or not the S-shaped propagation of new linguistic forms should be attributed to probabilistic weights in an I-language Grammar. I talked about the negation data set and initial tests of the Constant Rate Hypothesis at the conference *New Ways of Analyzing Variation (NWAV) 43* in Chicago in October 2014. I am most indebted to Anthony Kroch for invaluable comments and his encouragement. No less important was the University of Pennsylvania body of graduate students, who made the conference a very enjoyable event. I was afforded the opportunity to present this study as an invited speaker at the *Linguistics and English Language (LEL) Research Seminar* at the University of Manchester in November 2015. I would like to convey my thanks to George Walkden for the invitation, David Denison for comments on the additional variant *have got* and spotting incorrect examples, as well as the seminar audience for helpful feedback. The adverb data was independently presented in 2016 at the *Swiss Work in English Language and Linguistics 2016* in Fribourg as well as the *Symposium on the History of English Syntax 14* in Edinburgh. I would like to thank Marianne Hundt for pointing out potentially confusing terminology as well as the other participants of these conferences for stimulating discussions.

stative senses of concrete (*have a car*), but also abstract possession (*have an idea*), alienable (*have a house*) and inalienable ownership (*have blue eyes*), or descriptions of kinship (*have a brother*) as well as more dynamic senses of *have*, which may metaphorically express processes (*have a discussion*), function as opaque light verbs (*have a laugh*), or form lexicalized idioms (*have fun*).

The realization of possessive *have* in this sense changed substantially in certain syntactic contexts between the 19th and 21st centuries in standard American English. This is illustrated in the examples below. Example (1) exemplifies the conservative, example (2) the innovative, usage of possessive *have*.

(1) *conservative possessive* have
**Have we not** botanical gardens? **We have**, indeed, and much good they should do ...
Sir Samuel White Baker (1855) *Eight Years' Wanderings in Ceylon*

(2) *innovative possessive* have
"**Don't you have** a spare key?" "**I do**." He went into the house and returned with a large iron key ...
Frederick Ramsay (2008) *Stranger Room*

In (1) *not* negates the proposition directly, while in (2) negation requires *do*-support. Questions are formed by inverting *have* and the subject in (1), but by inserting the dummy auxiliary *do* in (2). Finally, (1) forms an elliptical answer with *have*, whereas (2), once again, uses the word *do*.

Thus, there are several contexts, such as negation, interrogatives and ellipsis, in which possessive *have* develops an innovative *do*-support feature. According to the Constant Rate Hypothesis, the linguistic innovation should spread at the same rate of change in each of these contexts. This, then, is the primary hypothesis to be tested in this chapter.

The chapter is structured as follows. The first part is theoretical in nature. I will develop a grammar fragment of Modern English that can generate the variation seen in (1) and (2) by employing the well-known category distinction between main and auxiliary verbs. This grammar can then be used to pinpoint formally the linguistic change in possessive *have*. Furthermore, the grammar model will also allow deducing the relevant contexts in which the underlying change in possessive *have* should surface at identical rates. The second part deals with the empirical material used to test the CRH. I will explain the search scripting principles according to which I queried the database used for the collection of the material, the *Corpus of Historical American English* (Davies 2010). I will then describe how the data was collected and analyzed for each linguistic context individually. In the third part, the CRH can finally be tested with the quantitative data thus obtained. A summary will complete the chapter.

## 2.2 Modeling the Change in Possessive *have*

In the first, theoretical part of this chapter, I will offer a formalization of the change in late Modern American English possessive *have*. Specifically, I will outline a grammar model of Modern English and use it to formally characterize the change in possessive *have* and to define discrete contexts in which the change should occur.

### 2.2.1 A Model of Modern English Clause Structure

The grammar fragment presented in this section includes at its core a distinction between main and auxiliary verbs. It follows in this respect standard accounts of Modern English clause structure, as taught frequently in elementary syntax courses, or as described in introductory syntax text books (e.g., Adger 2003, Radford 2004 etc.). A great deal of unnecessary detail has been omitted, and peripheral aspects of the model are dealt with only superficially. Thus, while the specific implementation of the theory may be unfamiliar, its fundamental facets should be relatively uncontentious.

#### 2.2.1.1 Auxiliaries and Full Lexical Verbs

Modern English has a well-known verbal distinction between auxiliaries and full lexical verbs. The first class consists of modals, like *may*, *will*, *must* etc., the dummy auxiliary *do*, perfect *have*, as well as passive, progressive and copula *be*. These verbs are illustrated in (3).

(3) *Illustration of auxiliaries*

a. *Modal 'may'*
Mary may study for her exam.

b. *Dummy 'do'*
My sister does not really play golf.

c. *Perfect 'have'*
John has visited Spain.

d. *Progressive 'be'*
You were spilling the beans too early.

Auxiliaries do not have a full argument structure but instead are feature bearing modifiers of the main verb. This is an oversimplification because copula *be* indisputably subcategorizes for a predicate, and there is evidence that some auxiliaries are in fact argument structure heads while others are not (Falk 2008). For the purposes of this discussion, however, I will ignore copula *be* and make the simplifying assumption that all other auxiliaries are merely morphological markers of verbal features, such as aspect, voice or modality. Auxiliaries are assumed to be of category `I`. This is exemplified in the sample lexicon entries in (4), where the letter after a word's graphemic form indicates its category. I utilize a verb form feature, `VFORM`, to implement a simple selection mechanism. The `VFORM` feature can take on the values infinitive, `inf`, past / passive participle, `pp`, present participle, `presp`, or finite verb, `fin`. Auxiliaries constrain the `VFORM` feature of their main verb. For example, the modal *may* in (4a) requires that the main verb should value its verb form feature as infinitive, $\texttt{VFORM}=_c\texttt{inf}$. Note that this formalization of selection is only rudimentary and extremely simplistic. Most importantly, it cannot account for multiple auxiliaries in the same clause. An adequate treatment of the topic would include either an implementation of morphological structure as a natural level for the representation of formal properties of verbs, like finiteness, the form of dependents etc. (Frank and Zaenen 2002), or use complex categories, such as `VP[prog]`, `VP[perf]` etc., instead of just one global `VP` category (Butt et al. 1999). However, I shall ignore all of these complexities here in order to keep the model simple. For expository purposes, the sample lexicon entries also include other features for each auxiliary, such as tense, aspect, modality and subject agreement, but they are not crucial to the overall argument.

(4)

a. *may* I $(\uparrow \texttt{VFORM})=_c \texttt{inf}$
$(\uparrow \texttt{TENSE})= \texttt{present}$
$(\uparrow \texttt{MODALITY})= +$

b. *does* I $(\uparrow \texttt{VFORM})=_c \texttt{inf}$
$(\uparrow \texttt{TENSE})= \texttt{present}$
$(\uparrow \texttt{SUBJ PERS})=_c 3$
$(\uparrow \texttt{SUBJ NUM})=_c \texttt{sg}$

c. *has* I $(\uparrow \texttt{VFORM})=_c \texttt{pp}$
$(\uparrow \texttt{TENSE})= \texttt{present}$
$(\uparrow \texttt{PERF})= +$
$(\uparrow \texttt{SUBJ PERS})=_c 3$
$(\uparrow \texttt{SUBJ NUM})=_c \texttt{sg}$

d. *were* I $(\uparrow \texttt{VFORM})=_c \texttt{presp}$
$(\uparrow \texttt{TENSE})= \texttt{past}$
$(\uparrow \texttt{PROG})= +$
$\{ (\uparrow \texttt{SUBJ PERS})=_c 2$
$(\uparrow \texttt{SUBJ NUM})=_c \texttt{sg}$
$\mid (\uparrow \texttt{SUBJ NUM})=_c \texttt{pl} \}$

The second class of verbs are full lexical verbs. They are characterized by their full argument structures, subcategorizing for entire grammatical functions, like subjects and objects. They can be finite or non-finite. Examples of lexical verbs are *want*, *play*, *finish* or *spill*, as shown in (5).

(5) *Illustration of full lexical verbs*

a. Mary wants to study for her exam.

b. My sister really plays golf.

c. John visited Spain.

d. You were spilling the beans too early.

Full lexical verbs are modeled as elements of category `V`. This is illustrated in the sample lexicon entries in (6), where, as before, the letter following the verb's graphemic form represents its category. I include the subcategorization frame for each verb as it would be required for parsing the sentences in (5). Every verb is typed for a particular finite or non-finite form by means of the `VFORM` feature. For example, the verb *wants* in (6a) is specified as a finite verb, `VFORM=fin`. I also show, again for expository purposes only, a few more lexical features and constraints for each entry, such as tense, subject agreement and control specifications.

(6) a. *wants* V (↑ PRED)= ‘WANT <(↑ SUBJ) (↑ XCOMP)>’
(↑ VFORM)= fin
(↑ SUBJ)= (↑ XCOMP SUBJ)
(↑ TENSE)= present
(↑ SUBJ PERS)$=_c$ 3
(↑ SUBJ NUM)$=_c$ sg

b. *plays* V (↑ PRED)= ‘PLAY <(↑ SUBJ) (↑ OBJ)>’
(↑ VFORM)= fin
(↑ TENSE)= present
(↑ SUBJ PERS)$=_c$ 3
(↑ SUBJ NUM)$=_c$ sg

c. *visited* V (↑ PRED)= ‘FINISH <(↑ SUBJ) (↑ OBJ)>’
(↑ VFORM)= fin
(↑ TENSE)= past

d. *spilling* V (↑ PRED)= ‘SPILL <(↑ SUBJ)>’
(↑ VFORM)= presp

A cluster of linguistic features separate the two classes of verbs. In terms of syntax, full lexical verbs, but not auxiliaries, require *do*-support for sentential negation, in inversion contexts, like direct questions, and for cases of VP-ellipsis. (These are three of the so-called NICE properties, Huddleston 1976; for a summary, see Huddleston and Pullum 2002: 92-112.) Further, full lexical verbs, but not auxiliaries, disallow certain VP-adjuncts in post-verbal position (Pollock 1989). There are also morpho-phonological differences between auxiliaries and full lexical verbs. The former but not the latter can be stressed directly for the purpose of emphatic polarity (the fourth NICE property), permit cliticization of negation, or form clitics or reduced forms.

In the following sections, I will propose a formalization of the syntactic facts listed above. I will not attempt to model the morpho-phonological characteristics because they are difficult to screen in naturally occurring corpus data and are thus not relevant for the ensuing corpus study on possessive *have*.

#### 2.2.1.2 The Clausal Hierarchy of Projections

My grammar model of English employs three standardly assumed projections, CP, IP and VP. The three projections select each other such that the first forms the highest, the last the lowest, element of the hierarchy. I will consider only finite matrix clauses, the start symbol for which is indicated as $\text{CP}_{MAT}$.

This section implements specific rules generating the hierarchy of projections for simple, declarative clauses. I will first consider three scenarios for the substitution of the $\text{CP}_{MAT}$ start node. In all of them, the restriction of $\text{CP}_{MAT}$ to finite clauses is formally achieved by requiring a TENSE feature to exist somewhere within its domain. Firstly, the specifier of $\text{CP}_{MAT}$ may remain unprojected, for example in ordinary subject-initial sentences, (7a). Secondly, $\text{CP}_{MAT}$'s specifier can host a fronted constituent, which I will assume to simply map to some underspecified discourse function, UDF. I will not formalize in any detail the actual information-structural import of this position (such as givenness, aboutness, familiarity etc.). If the category of the initial constituent is an adverb phrase, ADVP, the UDF must be a member of an adjunct set, ADJUNCT, either of the same clause or of an arbitrarily deeply embedded clausal complement, COMP or XCOMP. The head of the adjunct should neither be typed as a *wh*-element, ADV-TYPE≠wh, nor introduce a negative polarity feature, ↓ POL≠ neg. If these requirements are met, the clause type will be expected to be declarative, decl, (7b). Third, the initial constituent mapped onto UDF may also be realized as a determiner phrase, DP. In this case, the initial constituent realizes a governable grammatical function of the same or some arbitrarily deeply embedded clause. As before, if the UDF is neither a *wh*-pronoun, PRON-TYPE≠wh, nor a negative initial constituent, ↓ POL≠ neg, the clause is interpreted as declarative, (7c).

The rules in (7) have several limitations. They do not exhaust all possible categories for clause-initial constituents, such as CPs, ADJPs, VPs etc. Nor do the rules reflect all possible extraction paths. For example, there are clause-initial DPs that function not as arguments but as adjuncts. Finally, for brevity reasons, the rules use inelegant *wh*-type features instead of more economical complex categories (e.g., DP[wh]) as they might be utilized in industrial grammars. For the purpose of this chapter, however, the rules will suffice while remaining relatively uncomplicated.

(7) a. $$\mathtt{CP}_{MAT} \rightarrow \begin{array}[t]{c} \mathtt{C'}_{MAT} \\ \uparrow = \downarrow \\ (\downarrow \mathtt{TENSE}) \end{array}$$

b. $$\mathtt{CP}_{MAT} \rightarrow \begin{array}[t]{c} \mathtt{ADVP} \\ (\uparrow \mathtt{UDF}) = \downarrow \\ (\uparrow \mathtt{UDF}) \in (\uparrow \{\mathtt{XCOMP}|\mathtt{COMP}\}^* \ \mathtt{ADJUNCT}) \\ (\downarrow \mathtt{ADV\text{-}TYPE}) \neq \mathtt{wh} \\ (\downarrow \mathtt{POL}) \neq \mathtt{neg} \\ (\uparrow \mathtt{CLAUSE\text{-}TYPE}) =_{c} \mathtt{decl} \end{array} \qquad \begin{array}[t]{c} \mathtt{C'}_{MAT} \\ \uparrow = \downarrow \\ (\downarrow \mathtt{TENSE}) \end{array}$$

c. $$\mathtt{CP}_{MAT} \rightarrow \begin{array}[t]{c} \mathtt{DP} \\ (\uparrow \mathtt{UDF}) = \downarrow \\ (\uparrow \mathtt{UDF}) = (\uparrow \{\mathtt{XCOMP}|\mathtt{COMP}\}^* \ \mathtt{GF}) \\ (\downarrow \mathtt{PRON\text{-}TYPE}) \neq \mathtt{wh} \\ (\downarrow \mathtt{POL}) \neq \mathtt{neg} \\ (\uparrow \mathtt{CLAUSE\text{-}TYPE}) =_{c} \mathtt{decl} \end{array} \qquad \begin{array}[t]{c} \mathtt{C'}_{MAT} \\ \uparrow = \downarrow \\ (\downarrow \mathtt{TENSE}) \end{array}$$

Outside of inversion contexts, $\mathtt{CP}_{MAT}$ does not project a head, and it is necessarily typed as declarative, `decl`. The complement of $\mathtt{CP}_{MAT}$ is `IP`, (8).

(8) $$\mathtt{C'}_{MAT} \rightarrow \begin{array}[t]{c} \mathtt{IP} \\ \uparrow = \downarrow \\ (\uparrow \mathtt{CLAUSE\text{-}TYPE}) = \mathtt{decl} \end{array}$$

The specifier of `IP` is a subject (`SUBJ`) position in Modern English, (9).

(9) $$\mathtt{IP} \rightarrow \begin{array}[t]{c} \mathtt{DP} \\ (\uparrow \mathtt{SUBJ}) = \downarrow \end{array} \qquad \begin{array}[t]{c} \mathtt{I'} \\ \uparrow = \downarrow \end{array}$$

`IP` will project a head if the clause contains an element of category `I`, i.e., an auxiliary, in post-subject position. In this case, the clause type must be valued as declarative, (10a). Otherwise, the head of `IP` will not project at all, (10b). In either case, the complement of the `I`-projection is `VP`.

(10) a. $$\mathtt{I'} \rightarrow \begin{array}[t]{c} \mathtt{I} \\ \uparrow = \downarrow \\ (\uparrow \mathtt{CLAUSE\text{-}TYPE}) =_{c} \mathtt{decl} \end{array} \qquad \begin{array}[t]{c} \mathtt{VP} \\ \uparrow = \downarrow \end{array}$$

b. $$\mathtt{I'} \rightarrow \begin{array}[t]{c} \mathtt{VP} \\ \uparrow = \downarrow \end{array}$$

`VP` does not commonly occur with an overt specifier, (11a). The head of `VP` hosts full lexical verbs, both finite and non-finite. The complement of `V` can be realized in a number of different ways. However, for the sake of simplicity, I will assume that `VP`'s complement either maps on a nominal direct object, `OBJ`, or remains unprojected, as indicated with optionality brackets around the object, (11b).

(11) a. $$\mathtt{VP} \rightarrow \begin{array}[t]{c} \mathtt{V'} \\ \uparrow = \downarrow \end{array}$$

b. $$\mathtt{V'} \rightarrow \begin{array}[t]{c} \mathtt{V} \\ \uparrow = \downarrow \end{array} \qquad \begin{array}[t]{c} (\mathtt{DP}) \\ (\uparrow \mathtt{OBJ}) = \downarrow \end{array}$$

I will now illustrate how this model parses some simple declarative sentences of Modern English. Assume the following lexicon entries for the words *Mary*, *apples*, *does*, *is*, *eats*, *eat* and *eating*. The entry for *does* is repeated here from (4b). The form *eat* is ambiguous between a finite, non-third person singular verb and an infinitive. It therefore receives a disjunctive lexicon entry. Note that the auxiliaries *is* and *does* have the category label `I` while both finite *eats* and *eat* and non-finite *eating* and *(to) eat* are of category `V`.

(12) a. *Mary* N (↑ PRED)= ‘MARY’
(↑ PERS)= 3
(↑ NUM)= sg
(↑ NOUN-TYPE)= proper

b. *apples* N (↑ PRED)= ‘APPLE’
(↑ PERS)= 3
(↑ NUM)= pl
(↑ NOUN-TYPE)= common-count

c. *does* I (↑ VFORM)$=_c$ inf
(↑ TENSE)= present
(↑ SUBJ PERS)$=_c$ 3
(↑ SUBJ NUM)$=_c$ sg

d. *is* I (↑ VFORM)$=_c$ presp
(↑ TENSE)= present
(↑ PROG)= +
(↑ SUBJ PERS)$=_c$ 3
(↑ SUBJ NUM)$=_c$ sg

e. *eats* V (↑ PRED)= ‘EAT <(↑ SUBJ) (↑ OBJ)>’
(↑ VFORM)= fin
(↑ TENSE)= present
(↑ SUBJ PERS)$=_c$ 3
(↑ SUBJ NUM)$=_c$ sg

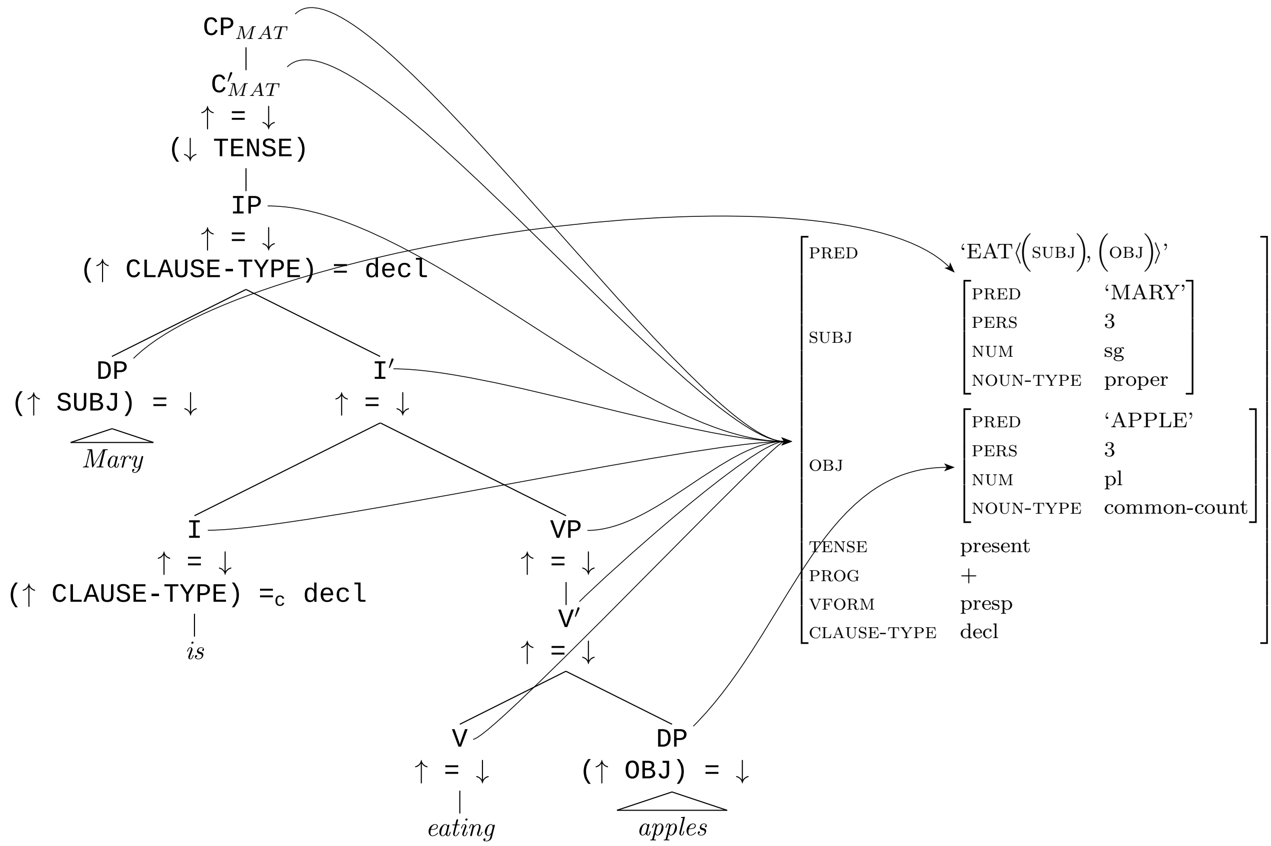
f. *eat* { V (↑ PRED)= ‘EAT <(↑ SUBJ) (↑ OBJ)>’
(↑ VFORM)= fin
(↑ TENSE)= present
¬ { (↑ SUBJ PERS)$=_c$ 3
(↑ SUBJ NUM)$=_c$ sg }
| V (↑ PRED)= ‘EAT <(↑ SUBJ) (↑ OBJ)>’
(↑ VFORM)= inf }

g. *eating* V (↑ PRED)= ‘EAT <(↑ SUBJ) (↑ OBJ)>’
(↑ VFORM)= presp

A fair number of example sentences can be parsed with the simple grammar model introduced in (7) to (11) and the toy lexicon in (12). For each example, there is a constituency tree generated by the grammar as well as an array of features that keeps track of the constraints imposed by the lexical and grammatical rules. Arrows are used to highlight the most important mappings between the two structures.

First, I will show how the grammar treats declarative sentences with an auxiliary and a non-finite main verb with the order ‘subject-aux-verb-object’, such as *Mary is eating apples*, (13). Here and in the subsequent examples, arrows highlight important mappings from c-structure to f-structure. In this first example, I include an arrow for every single node. Later, only pivotal parsing aspects will be signalized with arrows.

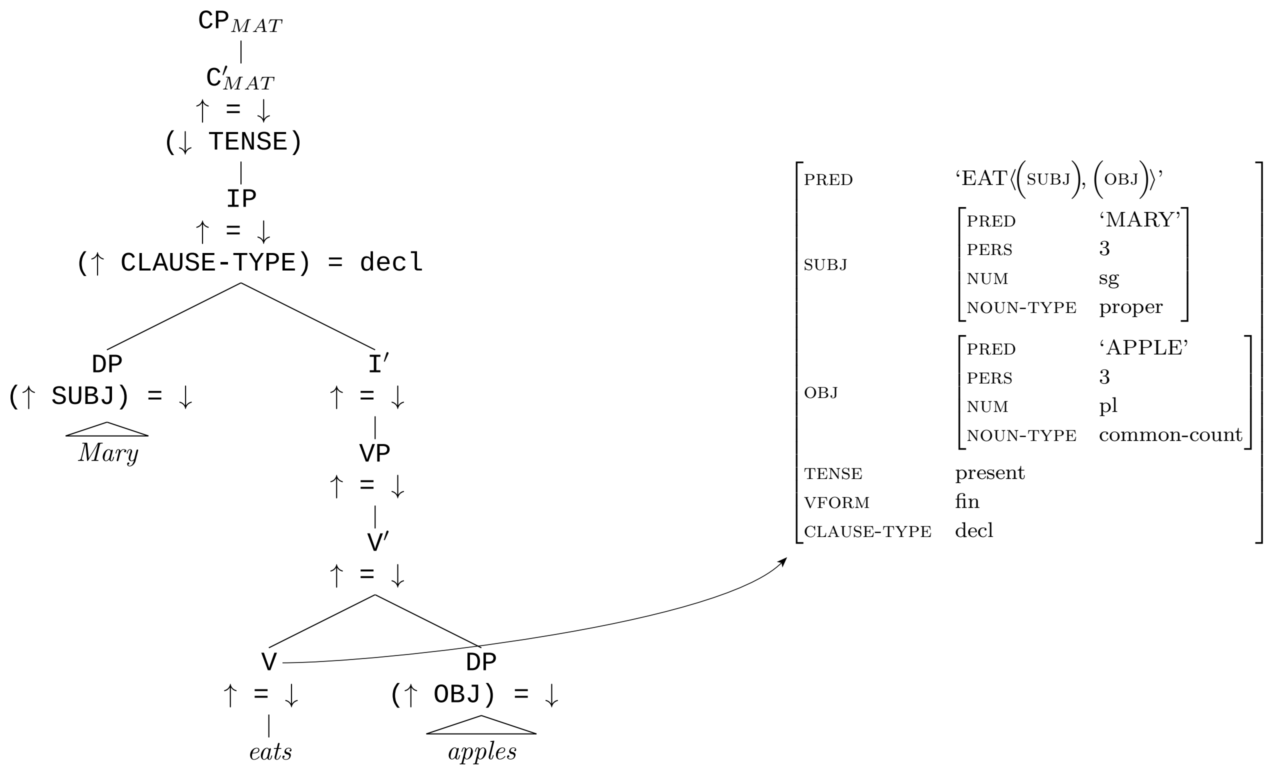
(13) Mary is eating apples.




The parser unifies all features that are introduced along the spine of the clause stretching from CP$_{MAT}$ over I to V. These include an underspecified TENSE feature inserted in the CP-layer. It is valued by the auxiliary *is* as `present` and hence the requirement that the matrix clause be finite is satisfied. Since the `CP`-layer is not complex but simply re-writes to `IP`, the whole structure is typed as a declarative clause, `CLAUSE-TYPE=decl`. The presence of this feature is also required by the head `I` in post-subject position. The auxiliary *is* introduces progressive aspect, `PROG=+`, and selects, by means of a `VFORM` feature, a main verb that is realized as a present participle. This constraint is satisfied by the existence of the full lexical verb *eating*, which values, as desired, the `VFORM` feature as present participle, `presp`. Further, the main verb *eating* subcategorizes for a subject and an object. The specifier of `IP` introduces the subject, the complement of `V` the object function, thus satisfying the principle of completeness, which states that all grammatical functions required by a predicate must be present. I do not explicitly analyze the internal constituency of the two grammatical functions but use instead triangles labeled `DP`. Finally, the auxiliary *is* requires a 3rd person singular subject. Since the subject is in fact specified accordingly, `PERS=3` and `NUM=sg`, the person and number features, too, unify successfully. Hence, all constraints imposed by the lexical and grammatical rules are satisfied in this structure, leading to one unique solution for the parse.

Next, the grammar can also handle sentences with a finite full lexical verb with the order 'subject-verb-object', as in *Mary eats apples*, (14).

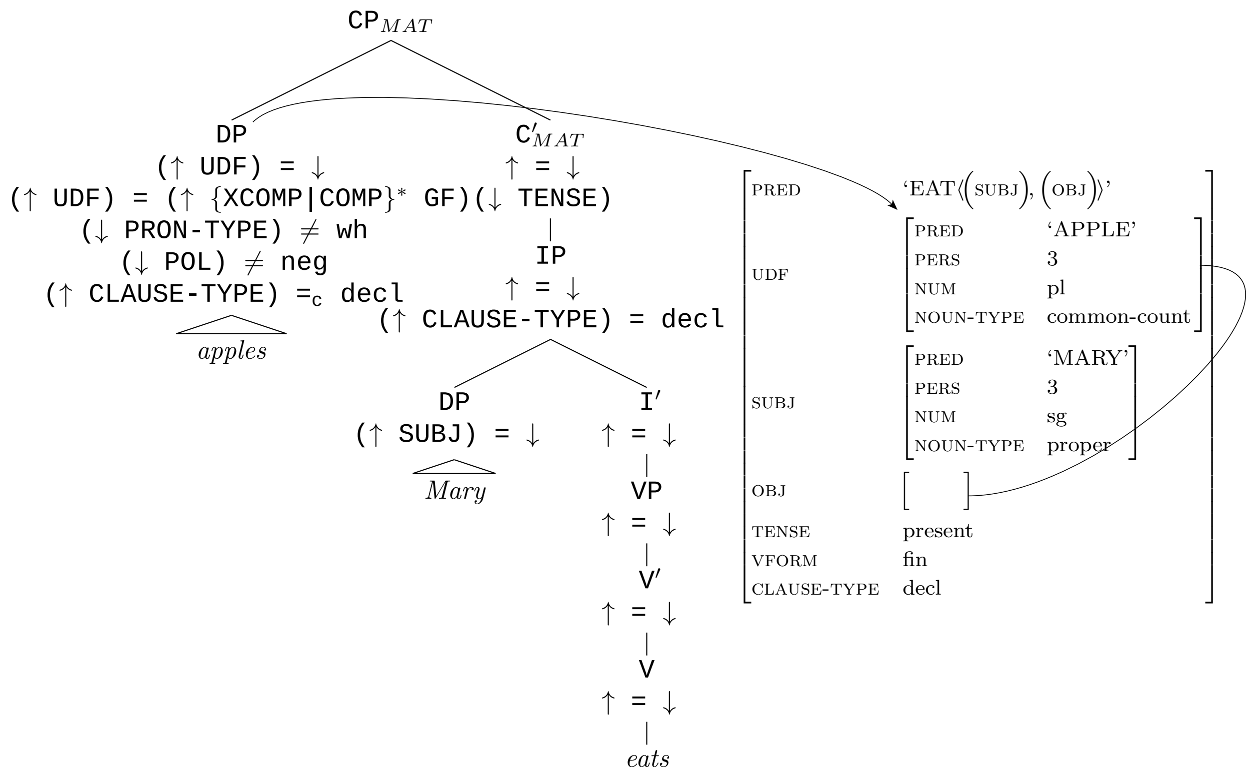
(14) Mary eats apples.




Whereas the tense feature was valued by an auxiliary in the last example, it now gets its value from the lexical verb *eats*, `TENSE=present`. This satisfies the constraint imposed by the matrix `CP` that the matrix clause must be finite. The main verb also introduces a `VFORM` feature. However, while the main verb was a present participle before, it is now valued as finite, `fin`, and is not involved in a selectional mechanism. Similarly agreement is now regulated by the lexical verb *eats*. It requires a 3rd person singular subject, and, as before, since the subject *Mary* is so specified, the agreement constraints are met. In all other respects, the parse is identical to the previous one. Hence, all requirements imposed on the structure by the lexical and grammatical rules are satisfied and the sentence can be parsed with one solution.

Declaratives with a fronted constituent are likewise parsable by the grammar developed. An example of such a structure is *apples, Mary eats*, (15).

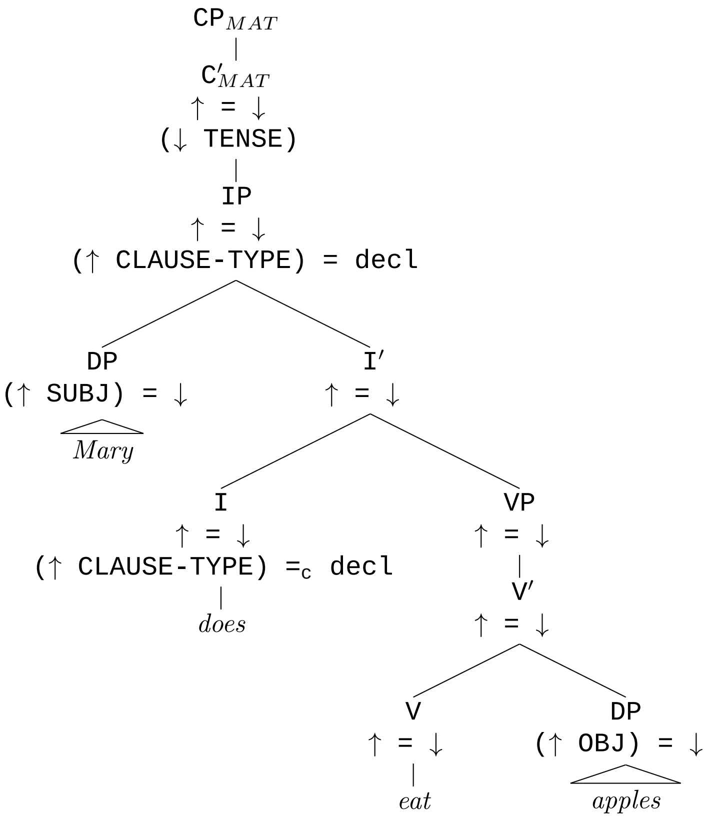
(15) Apples, Mary eats.




The fronted constituent *apples* is parsed as the specifier of $\texttt{CP}_{MAT}$. The grammar model cannot generate another analysis in a configuration where an initial `DP` immediately precedes the subject. In this position, the phrase must be mapped onto an underspecified discourse function, `UDF`. The constituent could be interpreted, for example, as a topic, but the specific information-structural reading is not relevant here. In principle, the grammar model allows for the initial, discourse-specific phrase to be associated with any governable grammatical function, `GF`. However, the set of grammatical functions subcategorized for by *eats* would, within the scope of the lexicon entries provided in (12), remain incomplete if it were interpreted as any element other than the object because the predicate would not be able to find its object function within the `IP`-domain *Mary eats*. As a consequence, the fronted constituent can only be understood as the object of the main verb. Specifically, the functionally uncertain constraint $(\uparrow \texttt{UDF})\texttt{=}(\uparrow \{\texttt{XCOMP|COMP}\}^{*}\ \texttt{GF})$ can only instantiate the constraint $(\uparrow \texttt{UDF})\texttt{=}(\uparrow \texttt{OBJ})$ in the current example. All other readings cannot be generated as grammatical parses by the model. The association between the discourse function, `UDF`, and the object function, `OBJ`, is represented by a line ranging between the two respective feature structures. The fact that $\texttt{C}'_{MAT}$ directly re-writes to `IP` signals a declarative interpretation of the sentence. Finally, the initial phrase, *apples*, does not include a `PRON-TYPE` or `POL` feature. Hence, the structure trivially satisfies the requirements that the fronted constituent should neither be typed as a *wh*-element, $(\downarrow \texttt{PRON-TYPE}){\neq}\texttt{wh}$, nor be a negative constituent, $(\downarrow \texttt{POL}){\neq}\texttt{neg}$. At the same time, the initial constituents requires that the clause be interpreted as declarative. This constraint is satisfied because the structure does in fact include a declarative clause type feature introduced by the `IP`-layer. Once again, all lexical and grammatical constraints are met in this example and a unique solution is produced. Constituent fronting of the kind illustrated here works identically in clauses with and without an auxiliary.

Note that the model can also analyze positive, declarative sentences with the dummy auxiliary *do*, such as *Mary does eat apples*, (16).

(16) Mary does eat apples.




$$\left[\begin{array}{ll}
\textsc{pred} & \text{`EAT}\langle(\textsc{subj}),(\textsc{obj})\rangle\text{'} \\
\textsc{subj} & \left[\begin{array}{ll} \textsc{pred} & \text{`MARY'} \\ \textsc{pers} & 3 \\ \textsc{num} & \text{sg} \\ \textsc{noun-type} & \text{proper} \end{array}\right] \\
\textsc{obj} & \left[\begin{array}{ll} \textsc{pred} & \text{`APPLE'} \\ \textsc{pers} & 3 \\ \textsc{num} & \text{pl} \\ \textsc{noun-type} & \text{common-count} \end{array}\right] \\
\textsc{tense} & \text{present} \\
\textsc{vform} & \text{inf} \\
\textsc{clause-type} & \text{decl}
\end{array}\right]$$

The auxiliary *does* values the tense of the clause as present, and constrains the subject to be 3$^{\text{rd}}$ person singular. It also requires that the main verb be realized as an infinitive. It does not make any additional feature contributions to the structure. In all other respects, the parse functions exactly as in (13), and has almost the same featural representation as *Mary eats apples* in (14).

Many syntacticians deem sentences such as (16) to be ungrammatical (e.g., Chomsky 1991: 62; Embick and Noyer 2001: 587 to name but a few), and might therefore criticize my model for being able to generate them. In this section, I will defend my assumption that all positive *do*-support sentences should be analyzable by the syntax. To do this, I will illustrate how alternative analyses function with simple but, I think, representative text book accounts, as in Haegeman & Guéron (1999: 528) or Radford (2004: 143).

A typical paradigm of acceptable and unacceptable positive *do*-support sentences is shown in (17).

(17) a. She **won** the battle (from: Haegeman and Guéron 1999: 528, ex. (57a))
b. * She **did win** the battle. (from: Haegeman and Guéron 1999: 528, ex. (57b))
c. Some people don't believe he won the race, but *he DID win it* (from: Radford 2004: 141, ex. (48f))

*Do*-support is viewed as a last resort mechanism, which means that dummy *do* is inserted "only when this is required to make the sentence grammatical" (Haegeman and Guéron 1999: 528). Formally, this is sometimes achieved by some economy matrix, for example counting and comparing the number of symbols in two derivations. If there is a grammatical sentence "with one symbol fewer" (ibid.) than some alternative, then this more economical derivation must be used. This redundancy explanation is supposed to rule out sentences like (17b). In order to allow sentences with emphatic *do*, such as (17c), it is usually assumed that the head of the clause bears either a phonological "stress feature" (ibid.) or an information-structurally sensitive "EMP(hasis) marker of some kind" for "contrastive structures" (Radford 2004: 143). *Do* is then assumed to be required as a host for this feature to be spelled out.

There are numerous, rather severe problems with this approach to positive *do*-support. Firstly, the last resort mechanism is actually formally ill-defined. As a context-sensitive transformation, *do*-insertion should specify explicitly and exhaustively the syntactic contexts in which it is appropriate. This would, in some way or other, boil down to richly specifying dummy *do* for each of its grammatical contexts, which could be considered inelegant. On the other hand, associating *do*-insertion with economy conditions enters formally unchartered territory. It is not clear how exactly such a rule should be formalized. Certainly, the comparison matrix suggested above is too simplistic. For example, it predicts that yes/no-questions, like *Do they play tennis?*, should be ungrammatical since they could be spelled out more economically without *do*-support, *They play tennis?*, with a rising question intonation. Thus, analyses of non-emphatic, positive *do* along these lines actually struggle to provide an explicit algorithm that rules out the supposedly offending structures.

Secondly, the account of grammatical emphatic *do* is problematic too. (i) Emphatic *do* could be supposed to spell out information-structurally relevant features, like contrast, as suggested by Radford. But there are several (e.g., Luzón Marco 1999), at least two distinct (Wilder 2013), pragmatic contexts in which positive *do* can be used felicitously. A syntactic account of licensing emphatic *do* would thus require a large, potentially untraceable number of *do* items with distinct syntactic features for each pragmatic context. (ii) Emphatic *do* could also be assumed to realize only a phonological feature, as proposed by Haegeman & Guéron. One could then get away with postulating only two kinds of *do* - one as a dummy auxiliary in grammatical contexts, another as a bearer of stress. However, it would now also be necessary to double the vocabulary items for every other auxiliary because they, too, can realize emphatic stress. In addition, this line of reasoning does not account for the fact that, outside of contrastive structures, positive *do* can be licensed by putting stress on another item, such as an emphatic adverb.[2]

(18) ... what we've learned in the last few days about Mohammad Abdulazeez is that he **does inDEED** appear to have suffered from depression ...
FOX News, 21 July 2015 "The O'Reilly Factor", John Roberts on "Who is Mohammad Youssuf Abdulazeez?"
http://video.foxnews.com/v/4366510778001/who-is-mohammad-youssuf-abdulazeez [0:54]

All of these problems can be avoided if positive *do* is treated as a complex phenomenon of the phonology-syntax-pragmatics interfaces which cannot be handled adequately with syntactic mechanisms alone. Thus, one should allow the syntax to freely generate positive *do*-support structures, as in (16), and filter out illegitimate sentences as violations of pragmatic or phonological constraints instead. Of course, to make this view convincing, an actual implementation of the phonology-syntax and syntax-pragmatics interfaces would be required, which, it goes without saying, is completely beyond the scope of this chapter.

However, the logic of this approach can at least be somewhat clarified with analogies to other illicit structures.

(19) What did Mary eat?
% It is Mary who ate apples.

The sentence in (19) is infelicitous. *Mary* was mentioned before, and was hence topicalized by the context, i.e., the preceding *wh*-question. But *it*-clefts put their isolated constituent into focus. This creates an information-structural contradiction. Furthermore, if an *it*-cleft is produced, it requires particular phonological patterns. For instance, they are unacceptable with main stress on *it* etc. Linguists would not normally assume that the *it*-cleft in (19) is syntactically ill-formed and should be starred.
The situation seems to me to be entirely parallel for positive *do*-support structures.

(20) What does Mary do?
% Well, Mary does eat apples.

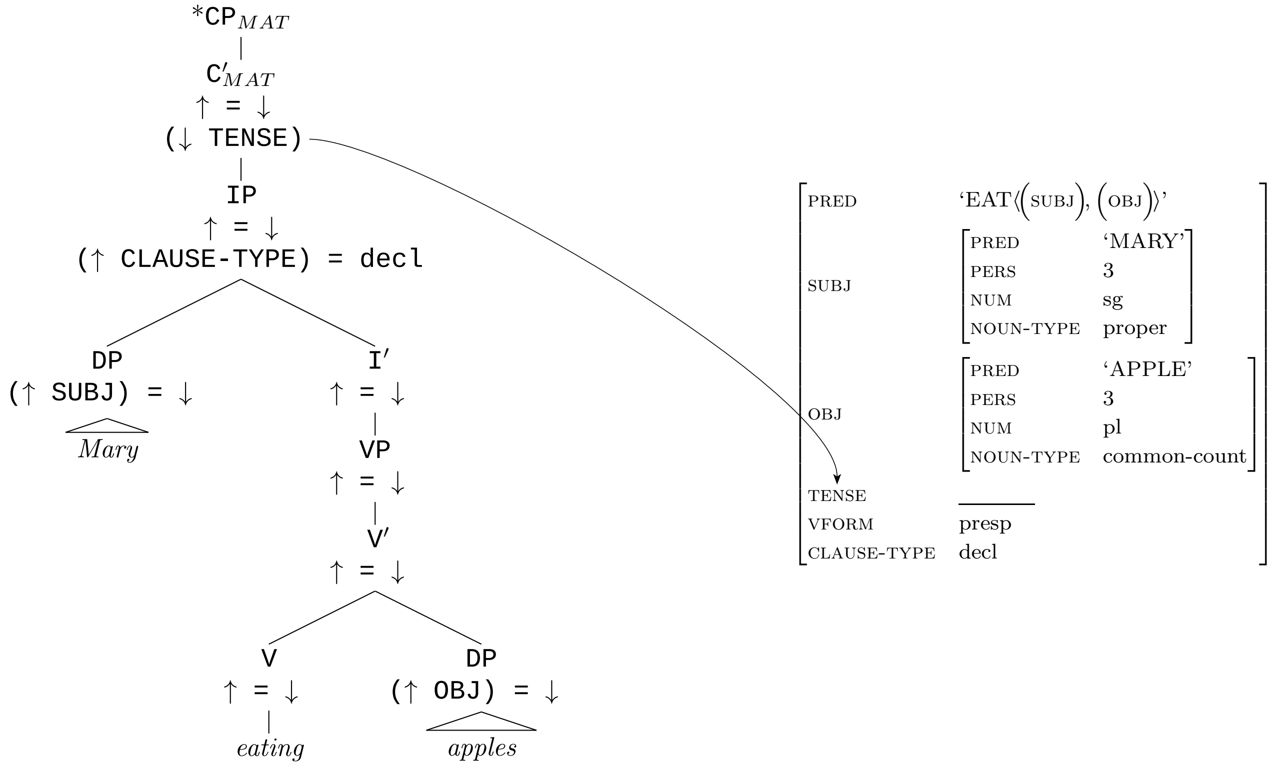
Just like the previous sentence, (20) is pragmatically odd. The reason seems to be, roughly, that there is no emotive or contrastive background that would make positive *do* contextually plausible. Further, again just as in the previous example, if a positive *do* sentence is formed, it is associated with permissible phonological contours, involving, for example, stress on the auxiliary etc. Despite these similarities, syntacticians usually insist that (20) is syntactically ungrammatical and must be starred. In actual fact, since the former example must not necessarily be considered syntactically ill-formed, the latter example probably should not be either.

In conclusion, it is a defensible aspect of my model that it seemingly overgenerates with respect to positive *do*-support structures. In fact, this may actually be a desirable corollary.

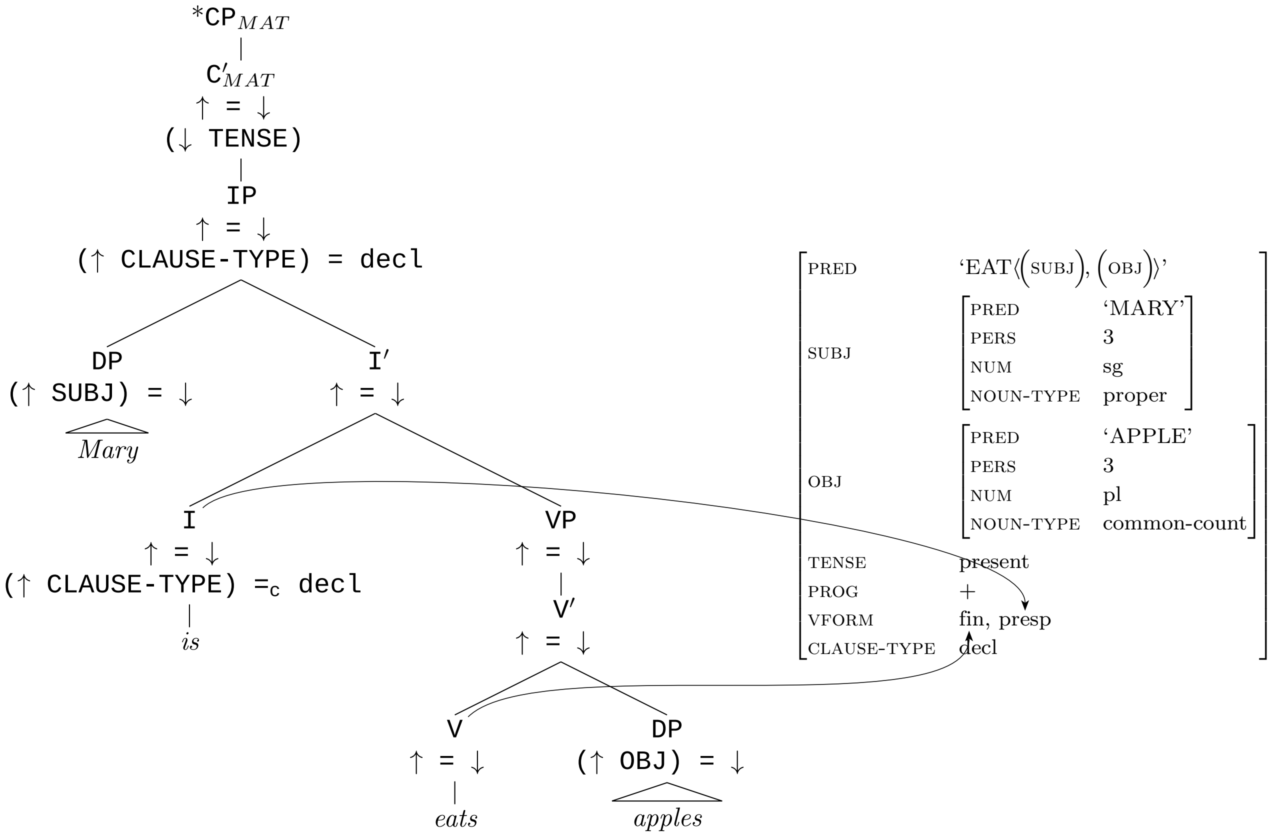
Finally, the grammar model can also identify certain ungrammatical sentences. In the following structures, a low line represents unsatisfied constraints, for instance an unvalued, i.e., missing, feature. Conflicting feature specifications are shown in the feature structure by listing the contradictory values next to each other. Examples of correctly rejected ungrammatical matrix clauses are **Mary eating apples*, (21), or **Mary is eats apples*, (22).

[2] I am grateful to Dara Jokilehto for suggesting this argument.

(21) * Mary eating apples.




(22) * Mary is eats apples.



In the first example, the finite matrix `CP` requires that the clause be tensed, but no tense feature is introduced anywhere in the structure. In particular, the present participle *eating* is not specified for tense. Consequently, the clause is not finite and the model cannot parse the string. The second example is ungrammatical because the progressive auxiliary *is* expects the main verb to be realized as a present participle. The `VFORM` features should thus be valued `presp`. However, the main verb *eats* is finite. It thus values the `VFORM` feature as `fin`. Since one feature attribute cannot take on two mutually exclusive values, a parse for the sentence cannot be produced.

These relatively detailed discussions serve to clarify the workings of the grammar at its simplest stage. In the next section, I will add a mechanism for sentential negation to the model.

#### 2.2.1.3 Negation

Sentential negation is one of the syntactic properties that distinguishes between auxiliaries and full lexical verbs in Modern English. The negator *not* is placed after auxiliaries (23a), but direct negation of full lexical verbs is ungrammatical (23b). Instead, they require *do*-supported negation, (23c), which is not a grammatical option for auxiliaries, (23d).

(23) a. Mary <u>may not</u> study for her exam.
b. * Mary <u>wants not</u> to study for her exam.
c. Mary <u>does not want</u> to study for her exam.
d. * Mary <u>does not may</u> study for her exam.

First, I will formalize the relatively simple case of negation contracted to auxiliaries, such as *isn't*, *won't*, *don't* etc. I adopt here a Lexicalist approach to the morphology-syntax interface (e.g., Bresnan 2001: 30, 92, 101; Spencer 2005: 81; Siddiqi 2014: 346), which is "the theoretical standpoint in modern generative linguistics according to which the processes that form complex words [...] are [in most cases] accounted for by a set of *Lexical Rules, independent of and different from* the syntactic rules of the grammar [...]. Such Lexical Rules are assumed to operate in a presyntactic component, the *Lexicon*" (Scalise and Guevara 2005: 147). Under such a view, auxiliary verbs with negative contraction are easy to handle because they can be assumed to be created by some morpho-lexical process that makes them available directly to the syntax. For example, regular, productive morphological processes can be analyzed with finite state transducers. They take a string as input (for example `d o n t`) and output a feature analysis string (for example `d o +Tense:Present +Pol:Neg`) (Beesley and Karttunen 2003), which could then be transformed into lexicon entries in a straightforward way. Forms of overgeneration (e.g. irregular plurals, **oxes*, but also irregular negative contraction, **willn't*) can be handled formally as well by overriding unwanted regular forms (ibid.: 300-309). More powerful formalisms are rarely needed to encode morphological processes. Alternatively, negatively contracted forms could simply be memorized idiosyncratically. The number of these words, about 20, is certainly small enough to be stored passively in the mental lexicon. In either scenario, auxiliaries with negative contraction will be accessible as complete lexical units (as in Kim and Sag 2002). There is also some evidence (Sadler 1998; Bender and Sag 2000) for the lexical integrity of similar forms, namely auxiliaries contracted to pronominal subjects, such as *you'll* or *I'd*. Example lexicon entries of negated auxiliaries are shown in (24). I record their category as `I+NEG`, which is mnemonic for 'auxiliary with negative contraction.' They introduce a polarity feature that is valued as negative, `POL=neg`. Otherwise, the negative auxiliaries are basically identical to their positive counterparts.

(24) a. *isn't* I+NEG $(\uparrow \text{VFORM}) =_c \text{presp}$
$(\uparrow \text{TENSE}) = \text{present}$
$(\uparrow \text{PROG}) = +$
$(\uparrow \text{SUBJ PERS}) =_c 3$
$(\uparrow \text{SUBJ NUM}) =_c \text{sg}$
$(\uparrow \text{POL}) = \text{neg}$

b. *won't* I+NEG $(\uparrow \text{VFORM}) = \text{inf}$
$(\uparrow \text{TENSE}) = \text{present}$
$(\uparrow \text{MODALITY}) = +$
$(\uparrow \text{POL}) = \text{neg}$

c. *don't* I+NEG $(\uparrow \text{VFORM}) =_c \text{inf}$
$(\uparrow \text{TENSE}) = \text{present}$
$\neg \{ (\uparrow \text{SUBJ PERS}) =_c 3$
$(\uparrow \text{SUBJ NUM}) =_c \text{sg} \}$
$(\uparrow \text{POL}) = \text{neg}$

d. *doesn't* I+NEG $(\uparrow \text{VFORM}) =_c \text{inf}$
$(\uparrow \text{TENSE}) = \text{present}$
$(\uparrow \text{SUBJ PERS}) =_c 3$
$(\uparrow \text{SUBJ NUM}) =_c \text{sg}$
$(\uparrow \text{POL}) = \text{neg}$

A grammatical rule is needed that re-writes `I'`, similar to the rule in (10a), but introduces the head of the projection as an element of category `I+NEG` instead of `I`. This alternative rule is shown in (25).

(25)
$$\texttt{I}' \rightarrow \quad \underset{(\uparrow\ \texttt{CLAUSE-TYPE}) =_c \texttt{decl}}{\underset{\uparrow = \downarrow}{\texttt{I+NEG}}} \quad \underset{\uparrow = \downarrow}{\texttt{VP}}$$

Next, I will discuss the more complex case in which the negator *not* functions as an independent word. One complicating factor in analyzing this lexical item is that it can be used in two different ways, namely for sentential negation and for constituent negation (for an overview, see for example, Zeijlstra 2015). Sentential negation inverts the polarity of a proposition such that it becomes true exactly under those circumstances when it would be false without the negator (even though scope interactions substantially complicate the matter). Constituent negation, on the other hand, does not invert the polarity of the entire clause but instead negates a smaller constituent, either a head (*a [not unpleasant] feeling*) or a phrase (*[not [very long ago]]*). It often occurs with a coordinated contrastive focus (*not Mary but Bill*, *not from London but from Paris* etc.). The Modern English employment of *do*-support in negation contexts, as illustrated in (23), has obliterated the ambiguity between these two types of negation in contexts where the negator *not* appears after a finite main verb. Such phrases can only involve constituent negation.

(26) She drinks not tea.
Sentential negation: *It is not the case that she drinks tea.
Constituent negation: It is the case that she drinks something, but not tea.

The ambiguity does, however, persist in all contexts in which the independent negator *not* appears between an auxiliary of category `I` and a main verb. The negation type is then usually difficult to ascertain, especially in naturally occurring, written data. I picked for the following two examples the auxiliary *could* because its distinct meanings can make the two negation types very salient. Sentential negation is likely when the auxiliary is interpreted in its past dynamic sense ('one was able to do something'), (27), but constituent negation is at least plausible with an irrealis deontic reading ('it would be good to do something'), (28). Some context is provided to induce one or the other negative reading. The (a) examples show that the linear order of the sentence elements is identical for both negation types and that the presence of a coordinated contrastive element is more likely with constituent negation. The examples in (b)-(f) demonstrate syntactic correlates of the two types of negation. Pseudo-clefts separate the negator and the element it scopes over in sentential negation but leave them together in constituent negation, (b). Tag questions typically reverse the polarity of a negated clause, but not of a negated sub-constituent of a clause, (c). Contracted negation can only be used for sentential negation, not for constituent negation, (d). Finally, if a clause involves sentential negation, it is modified by *either* and its `VP` pro-form is *neither*, but if merely a sub-constituent is negated, the polarity of the whole clause is not negative, and *too* and *so* are used respectively, (e), (f).

(27) *sentential negation*
The post office has lowered the price for a stamp every year for the last five years. But now there is a new law that forces the post office to increase prices. Here is what the post office was unable to do:
a. This year, the post office [could not [lower the price for a stamp ] ] (*but increase it).
b. What the post office could not do this year is lower the price for a stamp.
c. This year, the post office could not lower the price for a stamp, ?couldn't / could it?
d. This year, the post office couldn't lower the price for a stamp.
e. This year, the post office could not lower prices. And by the way, the supermarket couldn't, either.
f. This year, the post office could not lower prices. And by the way, neither could the supermarket.

(28) *constituent negation*
The post office has lowered the price for a stamp every year for the last five years. But now the post office is losing a lot of money. Here is how the post office could solve the problem:
a. This year, the post office [could [ [not lower] the price for a stamp ] ] (but increase it).
b. What the post office could do this year is not lower the price for a stamp.
c. The post office could not lower the price for a stamp, couldn't / *could it?
d. * This year, the post office couldn't lower the price for a stamp.
e. This year, the post office could not lower prices. And by the way, the supermarket could, too.
f. This year, the post office could not lower prices. And by the way, so could the supermarket.

The two uses of the negator can also co-occur.

(29) The post office was not going to lower the price for a stamp this year. But a new law forced the post office to. So now, the post office could not not lower the price for a stamp.

The syntax should thus be able to generate two distinct structures for the word order 'auxiliary - *not* - main verb,' one for sentential, another for constituent negation. For sentential negation, the negator should scope over the whole clause, valuing its polarity as negative. For constituent negation, the negator should adjoin to, and thus scope over, only the subconstituent it negates.

I am going to focus on sentential negation in the formal model that I am constructing here. I will not be concerned with formalizing constituent negation. The main reason for this restriction is that sentences with ambiguous negation cannot normally be disambiguated in corpus studies, such as the investigation of possessive *have* in the second part of this chapter, unless they also include a disambiguating syntactic device, like those in (27)/(28) (b)-(f). Further, there appears to be a strong bias towards interpreting occurrences of ambiguous negators as sentential negation. For example, sentences (28e) or (28f) seem distinctly odd when read out of context because, by default, *not* is assumed to scope over the whole sentence. Therefore it is reasonable to believe that the large majority of spontaneously produced negative sentences in a corpus will involve sentential and not constituent negation and to formally implement sentential negation only.

How, then, should the model incorporate negation with *not* in its function as a sentential negator? In principle, two constituency structures are conceivable. Either, the auxiliary and the negator form a constituent to the exclusion of VP, (30a). In derivational frameworks, this analysis could be likened to categorical movement of the negator to I. The result is a complex I-cluster. Alternatively, the negator and VP form a constituent to the exclusion of the auxiliary, (30b). In this case, the resulting larger constituent, NEGP, should exist in an empirically relevant sense.

(30) a. $[_{\mathbf{I}}$ aux *not*$]$ $[_{\mathbf{VP}}$ V ... $]$
b. aux $[_{\mathbf{NEGP}}$ *not* $[_{\mathbf{VP}}$ V ... $]$ $]$

The option in (30a) has been proposed explicitly for example by Otoguro (2006). He analyzes the negator *not* as a head-adjunct to I, (31a). The word *not* has the category label NEG and introduces a negation feature, NEG=+, (31b).

(31) a. $\texttt{I} \rightarrow \underset{\uparrow = \downarrow}{\texttt{I}} \quad \underset{\uparrow = \downarrow}{\texttt{(NEG)}}$
b. *not* NEG $(\uparrow$ NEG$)= +$

This formalization runs into two problems. Firstly, constituency tests point to the availability of a NEGP rather than a complex I-cluster. Admittedly, there are strong limits on the applicability of constituent tests for the two structures in (30). Most of the tests are inconclusive or cannot be used. For example, *do so* substitution will target the VP in either scenario, and the fact that the hypothesized NEGP cannot readily be used as a constituent in other structures, like fronting or clefts, could simply be due to the non-availability of the relevant phrase structure rules. However, there are a few tests that I believe can be employed successfully.
(i) Subject-auxiliary inversion is ungrammatical or at least severely degraded when the auxiliary co-occurs with independent *not*. In derivational terms, one could generalize that "*n't* moves along with the auxiliary to C [...]. The full form *not*, in contrast, does not move together with the auxiliary" (Repp 2009: 49). The following two examples illustrate.

(32) a. * Did not John leave? (from: Roberts 2010: 11, ex. (7))
b. * Would not he leave the city? (from: Kim 2001: 275, ex. (19a))

There are of course counterexamples to this generalization. Citing Quirk's *Comprehensive Grammar of English*, for instance, Haegeman (1995: 306, fn. 17) mentions questions such as *Is not history a social science?*. However, independent *not* in pre-subject position is quite formal and may simply be a printed equivalent of contracted negation (ibid.). It is exceedingly rare in casual conversation (e.g., Park 2008: 384-5). In my opinion, it would therefore be better if the grammar could not generate an auxiliary and free form *not* as one constituent.
(ii) Coordination facts also favor (30b) over (30a). Strings of the kind 'aux *not* and aux *not*' do not support the analysis of *not* as a head adjunct, but merely testify to the availability of VP for right-node raising, (33a). In contrast, right-node raised strings of the type '*not* VP' demonstrate fairly unequivocally the existence of a NEGP since they cannot be accounted for under a head-adjunction analysis of *not*, (33b). Furthermore, the string

'*not* VP' can be conjoined directly. One might claim that such structures involve constituent negation, but disambiguating syntactic devices, such as tag questions with inverted polarity, can at least reduce the likelihood of that option. Hence, they probably provide evidence for NEGP as well, (33c).

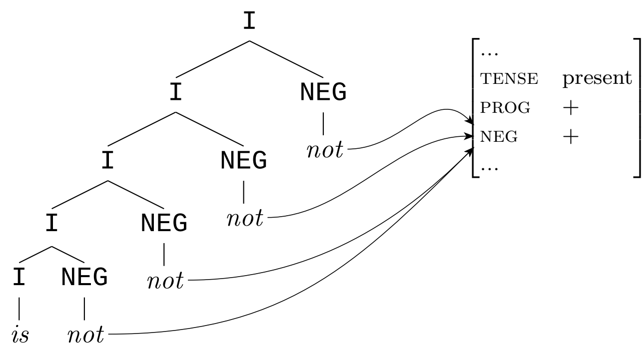
(33) a. Mary will not _, and should not _, [$_{\text{VP}}$ eat apples tonight].

b. Mary hates apples, and Bill is allergic to them.
So, Mary will _, and Bill should _, [$_{\text{NEGP}}$ not eat apples tonight].

c. Mary should [$_{\text{NEGP}}$ not eat apples tonight] and [$_{\text{NEGP}}$ not drink juice tomorrow], should she?

The second problem for Otoguro's analysis is that the phrase structure rule in (31a) is self-feeding and thus recursive. Multiple instances of the negator *not* all have the same feature contributions so that nothing will prevent their features from unifying. Therefore, the rule can generate structures with an infinite number of sentential negations. Further, an even number of negations should, if anything, reverse the negative interpretation of the clause in a language without negative concord, but, instead, such structures will value the polarity of the clause as negative as well. This is illustrated in (34). A grammar should be preferred in which there is only one position for sentential negation available.

(34)




The option in (30b), then, is probably more appropriate. Hence, I propose the following rules to implement sentential negation with free form *not* in Modern English:

(35) a. I′ → I NEGP

$$\begin{array}{ll} \uparrow = \downarrow & \uparrow = \downarrow \\ (\uparrow \text{ CLAUSE-TYPE}) =_c \text{decl} & \end{array}$$

b. NEGP → NEG VP

$$\begin{array}{ll} \uparrow = \downarrow & \uparrow = \downarrow \end{array}$$

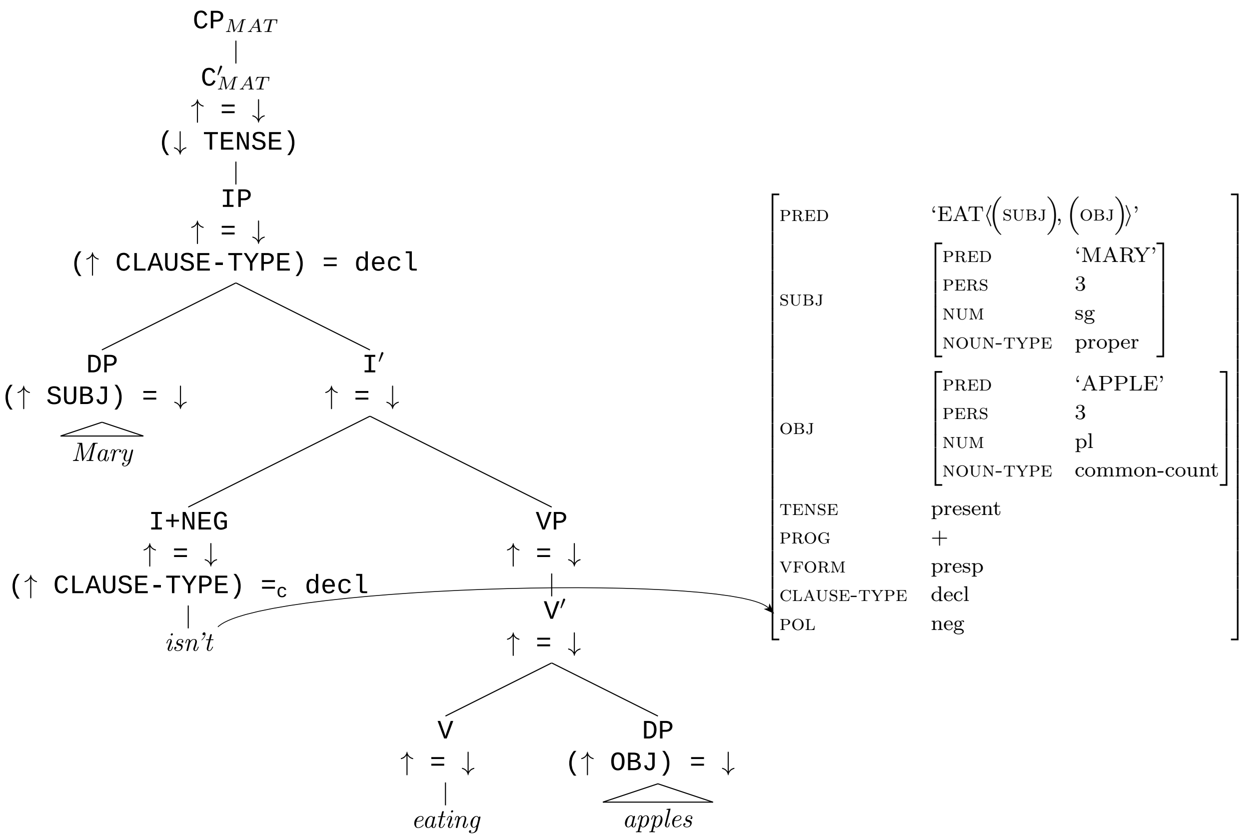
A dedicated projection introduces the independent negator *not*, NEGP, which is the complement of an auxiliary of category I. It takes VP as its complement. In this configuration, the clause is required to be declarative. Since NEGP dominates VP, it necessarily forms a constituent with it to the exclusion of I, as desired. I do not model a specifier position for NEGP because, as far as I can see, it is not needed. The lexicon entry for *not* is simply as in (36).

(36) *not* NEG (↑ POL)= neg

The feature specification for *not*, POL=neg, values the polarity of the structure it refers to as negative. The same feature can be formulated in any number of different ways, such as Otoguro's NEG=+, etc.

I will now illustrate how the model parses negative declarative clauses. First of all, the model can analyze sentences that contain an auxiliary with negative contraction, such as *Mary isn't eating apples*, (37).

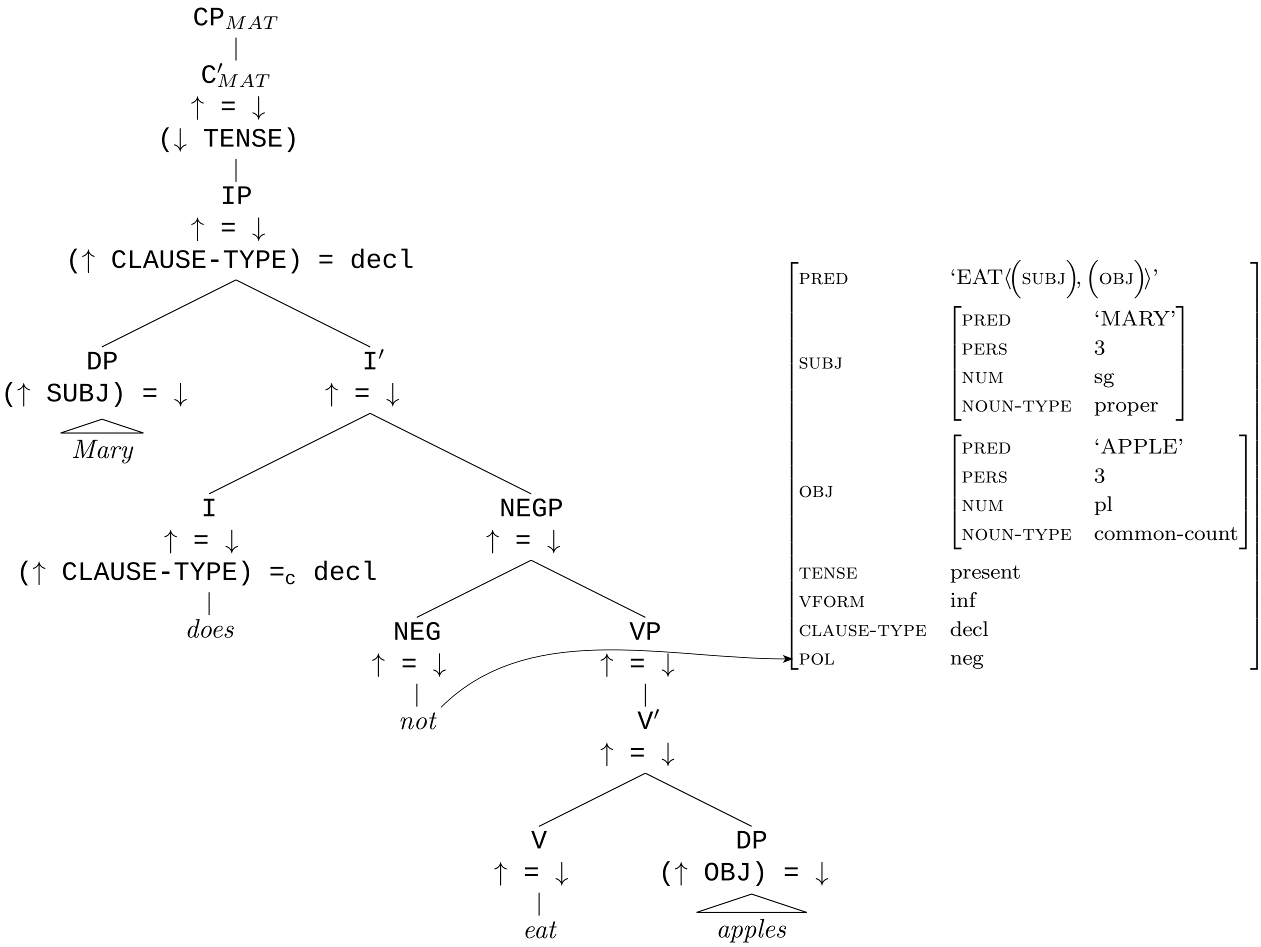
(37) Mary isn't eating apples.




Structures with negative contracted auxiliaries are headed by an element of category `I+NEG`, not by an auxiliary of category `I`. Lexical items of category `I+NEG`, such as *isn't* in the above example, value the polarity of the clause they head as negative, `POL=neg` (see the lexicon entry for *isn't* in (24a)). They are introduced into the structure by combining with `VP` (see the phrase structure rule in (25)). In all other respects, sentence (37) is analyzed exactly like its positive counterpart, *Mary is eating apples*, in (13). Hence, the model generates one unique solution for declarative clauses involving a negative contracted auxiliary.

Secondly, the grammar can analyze sentences with free form *not*, as in *Mary does not eat apples*, (38).

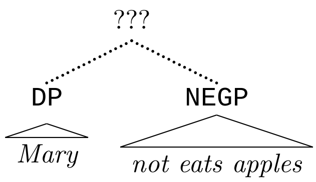
(38) Mary does not eat apples.




The independent negator *not* values the polarity of the clause as negative, POL=neg (see lexicon entry for *not* in (36)). The negator is of category NEG, which heads a NEGP and takes VP as its complement (see phrase structure rule in (35b)). The NEGP, in turn, co-occurs with an element of category I as its left sister. The auxiliary used in the current example is the dummy auxiliary *does*. It makes almost no featural contribution to the overall structure, but merely introduces present tense, satisfying the constraint that there must be a TENSE feature, requires a subject specified for 3$^{rd}$ person singular, such as *Mary*, and selects a main verb with an infinitival VFORM feature, such as *eat* (see lexicon entries for *does* in (4b) / (12c)). Moreover, the combination of I and NEGP constrains the clause to be declarative (see phrase structure rule in (35a)). Here, this requirement is met by virtue of the fact that the CP$_{MAT}$-layer is empty and directly re-writes to IP, which introduces the feature CLAUSE-TYPE=decl (see phrase structure rule in (8)). Thus, the structure satisfies all lexical and grammatical constraints imposed on it, and one unique parse for the string can be produced.

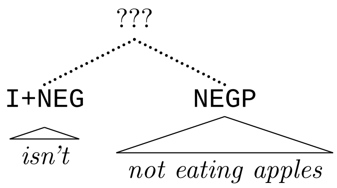
The grammar correctly rejects certain ungrammatical sentences with negation. Most importantly, it cannot generate clauses including a negated finite verb without *do*-support, such as **Mary not eats apples.* This is explained in a straightforward way by the non-existence of a licensing combinatory rule. With the rules discussed so far, the sentential negator *not* can be added to the structure only as the complement of an I-head, which must therefore necessarily be present. A phrase structure rule to combine the subject directly with NEGP is not available, (39).

(39) * Mary not eats apples.



The same mechanism will also preclude the co-occurrence of both a negative contracted auxiliary and free form *not* in the same clause. An example of such a sentence is **Mary isn't not eating apples*, intended as a clause with one sentential negation, 'It is not the case that Mary is eating apples.' In terms of feature unification, such structures are theoretically permissible because both `NEG+I` and `NEG` introduce the same negation feature, `POL=neg`, which would therefore not induce a mismatch. However, just as in the last example, these sentences cannot be generated simply because a suitable combinatory rule is lacking in the grammar. There is no phrase structure rule available that could directly combine the categories `I+NEG` and `NEGP` with each other, (40).

(40) * Mary isn't not eating apples.
intended: 'It is not the case that Mary is eating apples' (one sentential negation)




This concludes my discussion of negation in Modern English. In the subsequent section, I will extend the model to cover various contexts of subject-auxiliary inversion.

#### 2.2.1.4 Inversion

Subject-auxiliary inversion is another syntactic environment that displays a marked contrast between auxiliaries and full lexical verbs in Modern English. Inversion is required, for instance, to form ordinary, direct, non-subject questions. Auxiliaries invert directly with the subject, (41a), whereas full lexical verbs cannot do so, (41b). Instead, *do*-support is used to form direct questions in the absence of another auxiliary, (41c), but cannot be employed if another auxiliary is present, (41d). In the examples below, auxiliaries are underlined and the subject is highlighted in boldface.

(41)
a. Has **my sister** really played golf?
b. * Plays **my sister** really golf?
c. Does **my sister** really play golf?
d. * Does **my sister** really have played golf?

There are various contexts in which subject-auxiliary inversion is employed. Firstly, it is used in yes/no-questions, as in the examples above. The second context consists of questions that involve an initial non-subject *wh*-constituent. Both yes/no- and *wh*-questions may be negative, either by inverting a negatively contracted auxiliary, or by putting an independent negator after the subject. Thirdly, subject-auxiliary inversion can occur outside of interrogative clauses if a negative initial constituent is present. These are the only cases of subject-auxiliary inversion that I will include in my formal grammar model. I will ignore other, more marginal contexts, such as restrictive initial constituents (*only gradually do we make progress*), counterfactuals, which are largely restricted to *had* and *should* (*Had she hurried up, she would have arrived by now*) or wishes and curses, which usually occur with *may* (*May she rot in hell*) (e.g., Goldberg 2006: 166-182). These minor inversion contexts do not productively occur with possessive *have* and can therefore be ignored here.

I will first implement yes/no-questions. Their generation requires a new re-write rule for $\texttt{C}'_{MAT}$ that combines an overt `C`-head with `IP`. In this configuration, the clause must be interrogative, `CLAUSE-TYPE=interrog`, (42). The projection of a `C`-head and the introduction of an interrogative clause type feature differentiates this rule from its alternative in (8).

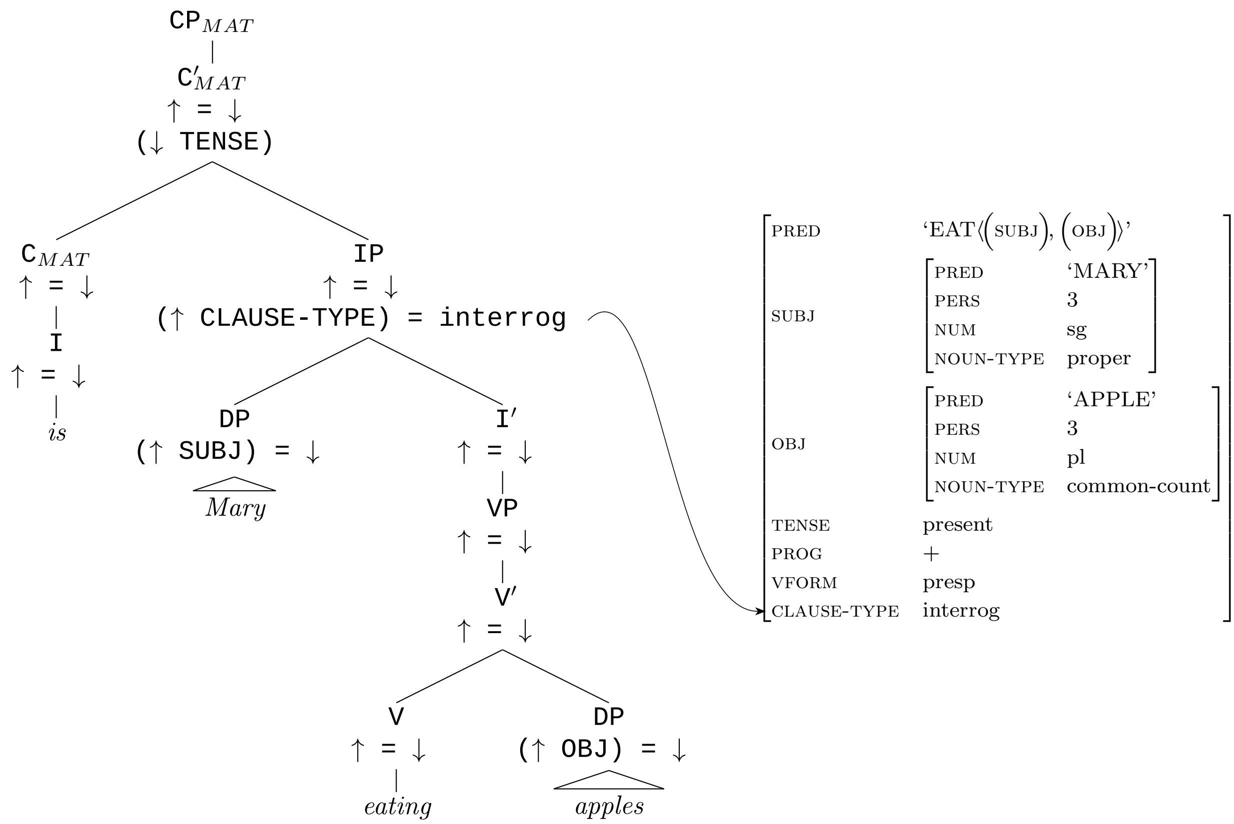
(42)
$$\texttt{C}'_{MAT} \rightarrow \begin{array}{c} \texttt{C}_{MAT} \\ \uparrow = \downarrow \\ \phantom{x} \end{array} \quad \begin{array}{c} \texttt{IP} \\ \uparrow = \downarrow \\ (\uparrow \texttt{ CLAUSE-TYPE}) = \texttt{interrog} \end{array}$$

The head of a finite matrix clause, $\texttt{C}_{MAT}$, can be re-written as a finite verb of category `I`, i.e., an auxiliary, (43). This process is often referred to as 'I-to-C movement' in derivational frameworks.

(43)
$$\texttt{C}_{MAT} \rightarrow \begin{array}{c} \texttt{I} \\ \uparrow = \downarrow \end{array}$$

These two additional rules suffice to successfully parse yes/no-questions. For example, the sentence *Is Mary eating apples?* is analyzed as in (44).

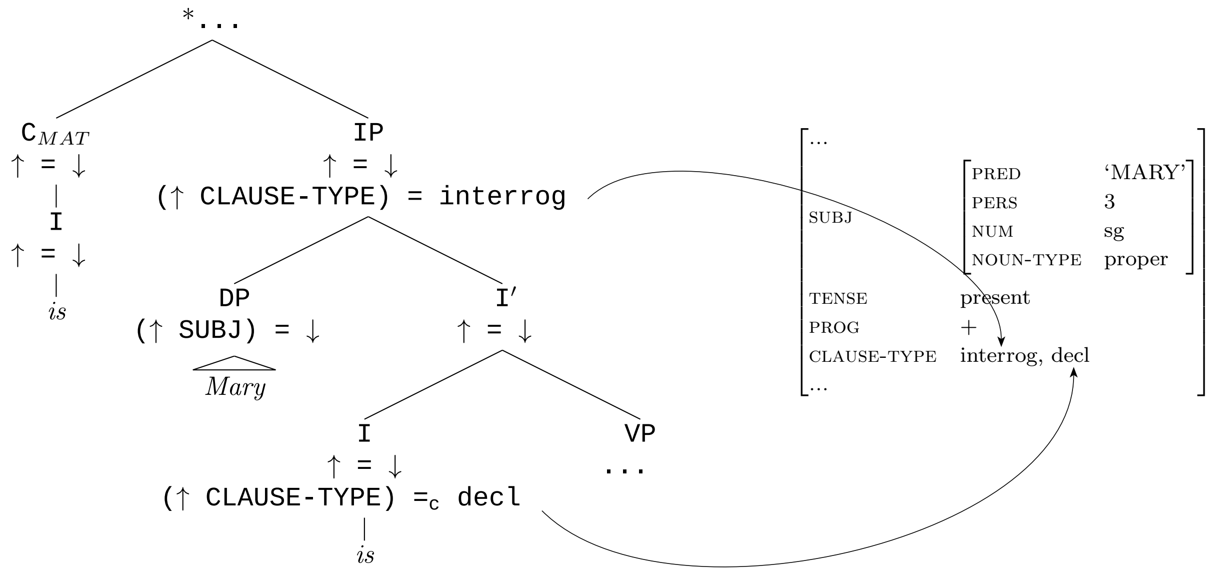
(44) Is Mary eating apples?




The parse must utilize the phrase structure rule in (42) to re-write $\mathtt{C'}_{MAT}$ to $\mathtt{C}_{MAT}$ and `IP`. Otherwise, it would be impossible to arrive at a hierarchical structure that places the finite verb in front of the subject. The application of this rule means that the clause must be interpreted as interrogative, `CLAUSE-TYPE=interrog`. The finite matrix `C`-head can only be re-written as an auxiliary of category `I`, as defined in rule (43). As before, the finite verb, *is*, satisfies the finiteness requirement by tensing the clause, controls subject verb agreement, introduces progressive aspect and selects a present participle as the main verb. *Mary* and *apples* are the subject and object functions of the clause respectively. The model generates exactly one solution for parsing the string under investigation.

Certain ungrammatical sentences are correctly rejected by the model as a consequence of the difference between interrogative and declarative clause type features. An example of such a sentence is **Is Mary is eating apples?*. In principle, the model permits the double occurrence of the auxiliary *is* since the features of both occurrences are identical and should thus be able to unify. In derivational frameworks, too, the structure may not be considered absurd a priori. Under a copy theory of movement (e.g., Chomsky 1995: 202-7, 251-3), it could result from multiple spell-out of both the auxiliary's higher instance and its lower copy. This option is sometimes assumed to be banned not by the syntax per se but by certain phonological conditions (e.g., Bošković and Nunes 2007). Under my proposal, the sentence cannot be generated because the pre-subject auxiliary (*is Mary...*) requires that the clause be typed interrogative, whereas an auxiliary in post-subject position (*Mary is...*) expects the presence of a declarative clause type feature. The feature mismatch induces ungrammaticality, as sketched in (45). Other forms of multiple auxiliaries will be banned for the same reason.

(45) * Is Mary is eating apples




Next, I will turn to *wh*-questions. Their generation requires the extension of the re-write rules for CP$_{MAT}$ in (7). As an illustration, I will present two rules, allowing a clause-initial adjunct ADVP or argument DP to be mapped onto a clause-initial discourse function. These rules check if the initial constituent is typed as a *wh*-phrase, and, if it is so typed, require that the clause receive an interrogative interpretation. Here, the initial *wh*-element can be unified with a local function or with a function of an embedded complement. The exact extraction paths have been discussed in some detail, for instance, by Dalrymple (2001: 405-8). The resulting rules are shown in (46).

(46) a. CP$_{MAT}$ → ADVP C′$_{MAT}$

ADVP: (↑ UDF) = ↓
(↑ UDF) ∈ (↑ {XCOMP|COMP}$^*$ ADJUNCT)
(↓ ADV-TYPE) =$_c$ wh
(↑ CLAUSE-TYPE) =$_c$ interrog

C′$_{MAT}$: ↑ = ↓
(↓ TENSE)

b. CP$_{MAT}$ → DP C′$_{MAT}$

DP: (↑ UDF) = ↓
(↑ UDF) = (↑ {XCOMP|COMP}$^*$ GF)
{(↓ PRON-TYPE) =$_c$ wh | (↓ D-TYPE) =$_c$ wh }
(↑ CLAUSE-TYPE) =$_c$ interrog

C′$_{MAT}$: ↑ = ↓
(↓ TENSE)

Analyzing *wh*-questions also requires, of course, the availability of some *wh*-items. I will consider the words *why* and *what* as examples. The first item is an adverb and carries a *wh*-type feature, ADV-TYPE=wh. It is assigned the lexicon entry in (47).

(47) *why* ADV (↑ PRED)= 'WHY'
(↑ ADV-TYPE)= wh

The *wh*- item *what* has two uses. It can either be a pronoun heading its own function with a PRED feature (as in *What is this?*). It can also be a *wh*-determiner without a PRED feature co-occurring with a nominal complement (as in *What town is this?*). The lexical item is typed as a *wh*-item in both uses, PRON-TYPE=wh, D-TYPE=wh. In order to differentiate *what* from *which*, I include a restrictiveness feature that is valued negatively for *what* and would be inversely specified for *which*, RESTR=-. *What* usually denotes an unrestricted set of many, potentially infinite possibilities whereas *which* commonly denotes a contextually restricted set of alternatives.

(48) a. <u>What towns</u> are we going to visit?
(=set of all towns, e.g., the speaker does not know the area, RESTR=-)

b. <u>Which towns</u> are we going to visit?
(=set of contextually relevant towns, e.g., the speaker knows the area, RESTR=+)

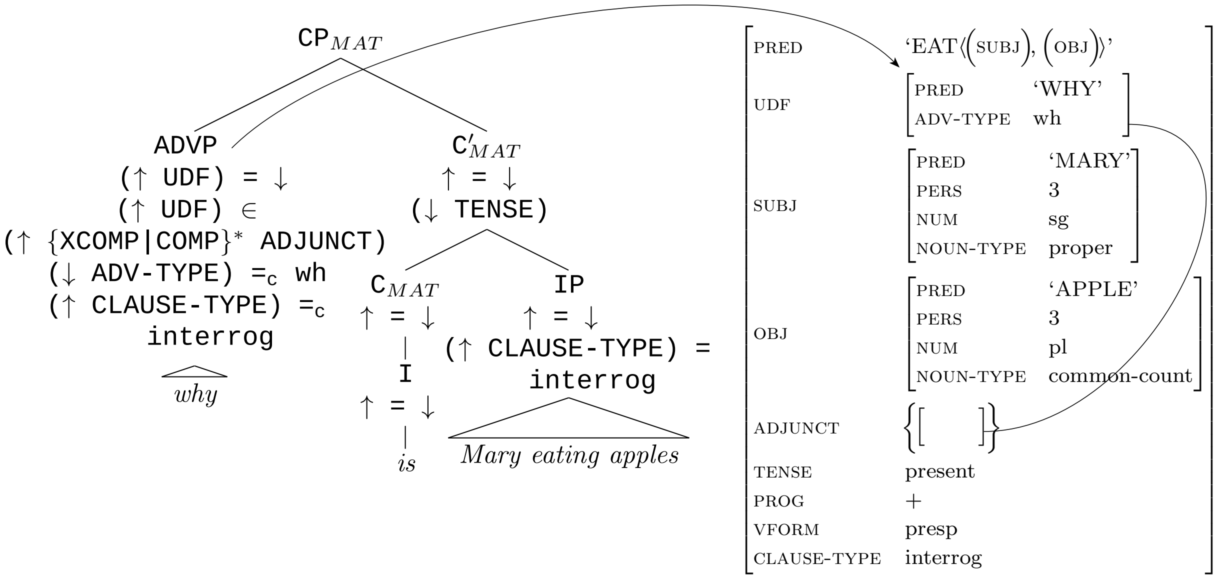
The restrictiveness feature is merely a syntactically irrelevant, mnemonic shorthand for some semantic computation performed on the f-structure that contains it. The lexicon entry for *what* is shown in (49).

(49) *what* D { (↑ PRED)= 'PRO'
(↑ PRON-TYPE)= wh
(↑ RESTR)= -
| (↑ D-TYPE)= wh
(↑ RESTR)= - }

Other *wh*-items will have similar lexicon entries in that they will all share the property of typing their function as a *wh*-phrase by means of a TYPE feature. As the above examples show, their exact specifications may be relatively diverse, realizing different syntactic categories, occurring with or without a PRED feature, or bearing additional semantic features such as RESTR etc.

The model can now also parse *wh*-questions, such as *Why is Mary eating apples?*, (50).

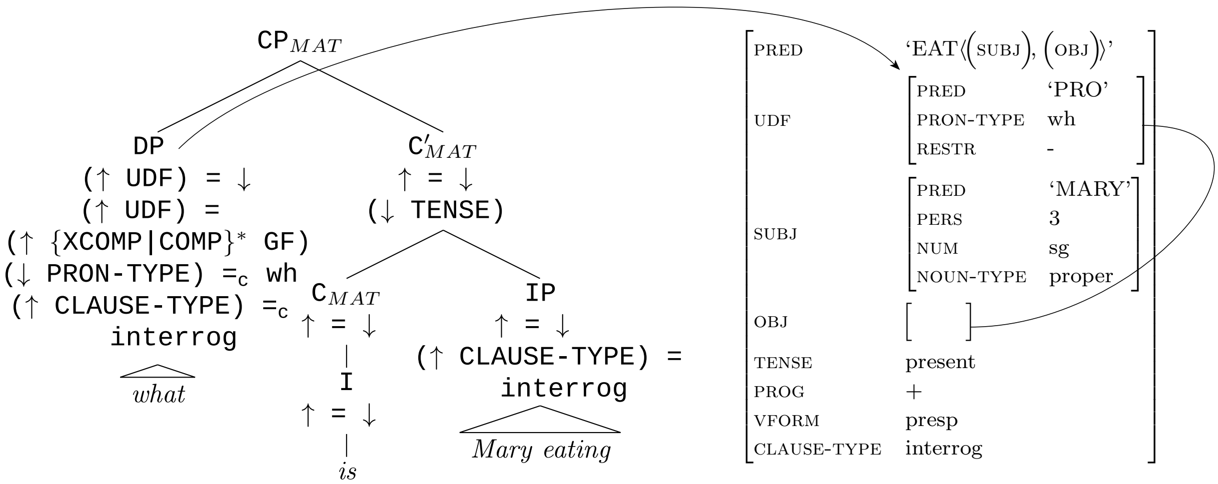
(50) Why is Mary eating apples?




The clause-initial adverb phrase, *why*, bears a type feature, ADV-TYPE=wh, and is thus mapped onto an underspecified discourse function of an interrogative clause, evidently some kind of question focus, by the rule in (46a). The required interrogative clause type feature is introduced by the subject-auxiliary inversion rule in (42). The discourse function is unified with a member of an adjunct set. This is represented by a line ranging between the discourse function and an adjunct in the feature array. In all other regards, the analysis functions as before. All constraints are satisfied and the model arrives at one unique solution for the parse.

Object questions are handled similarly. For example, *What is Mary eating?* is analyzed as in (51).

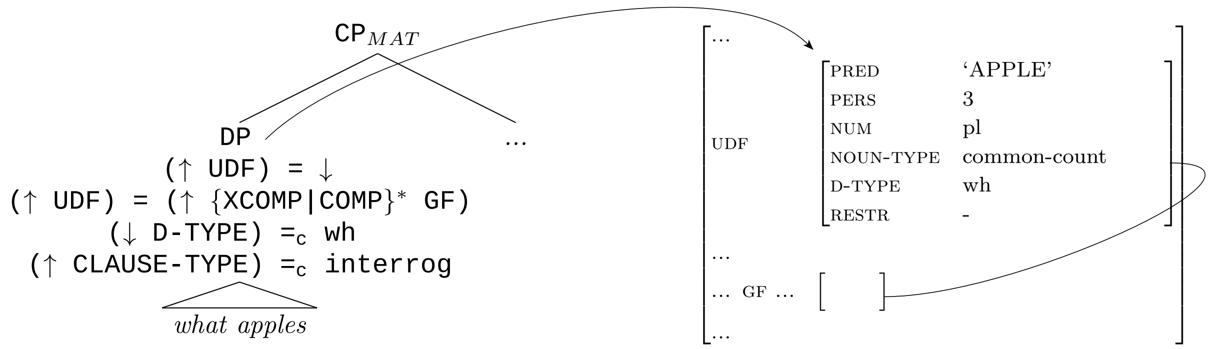
(51) What is Mary eating?



The clause-initial DP, *what*, functions as an underspecified discourse function, as regulated by the rule in (46b). It must also correspond to a governable grammatical function, GF. As for the sentence *Apples, Mary eats* in (15) above, only the instantiation of the constraint (↑ UDF)=(↑ {XCOMP|COMP}$^*$ GF) as (↑ UDF)=(↑ OBJ) will result in a coherent and complete functional structure and must therefore be selected. Unification of those two functions is represented with a connecting line in the feature array. The discourse function must be headed by a PRED. As a consequence, *what* must be parsed with the lexicon entry as a pronoun and not as a *wh*-determiner in (49). Since *what* is typed as a *wh*-pronoun, PRON-TYPE=wh, the clause must express interrogative force. This requirement is met by the introduction of an interrog feature by the subject-auxiliary inversion rule in (42). All constraints are satisfied and the string is parsed with one solution.

Object questions with more complex *wh*-consituents are parsed in the same way. For example, the sentence *What apples is Mary eating?* is different from the last example only in terms of the nature of its initial phrase, as sketched in (52).

(52) What apples is Mary eating?




The initial discourse function must be headed by one and only one PRED feature. Therefore, the *wh*-determiner interpretation of *what* must be selected from the lexicon entry in (49). The features of *apples* and *what* can then unify, as shown in the feature structure of UDF above. The parse can then proceed as before.

The model correctly rejects direct *wh*-questions without subject-auxiliary inversion.

(53) * What Mary is eating? (as a matrix, non-echo question)

The reason is that the initial *wh*-item expects the presence of interrogative force by rule (46b), but instead a declarative feature is introduced in the absence of subject-auxiliary inversion by rule (8). The contradictory feature specification results in ungrammaticality.

I will now discuss the interaction between inversion and sentential negation. Firstly, auxiliaries with contracted negation in front of subjects can successfully be parsed with only one slight modification of the model. The head of $CP_{MAT}$ must simply be allowed to re-write to a node of category I+NEG as well.

(54) $C_{MAT}$ → I+NEG
↑ = ↓

The grammar can now analyze inversion with negative auxiliaries, as in (55).

(55) [$_{C_{MAT}}$ [$_{\textbf{I+NEG}}$ Isn't ] ] Mary eating apples?

The parse works exactly as it did for *Is Mary eating apples?* in (44) except that the negatively contracted auxiliary will additionally introduce a negative polarity feature.

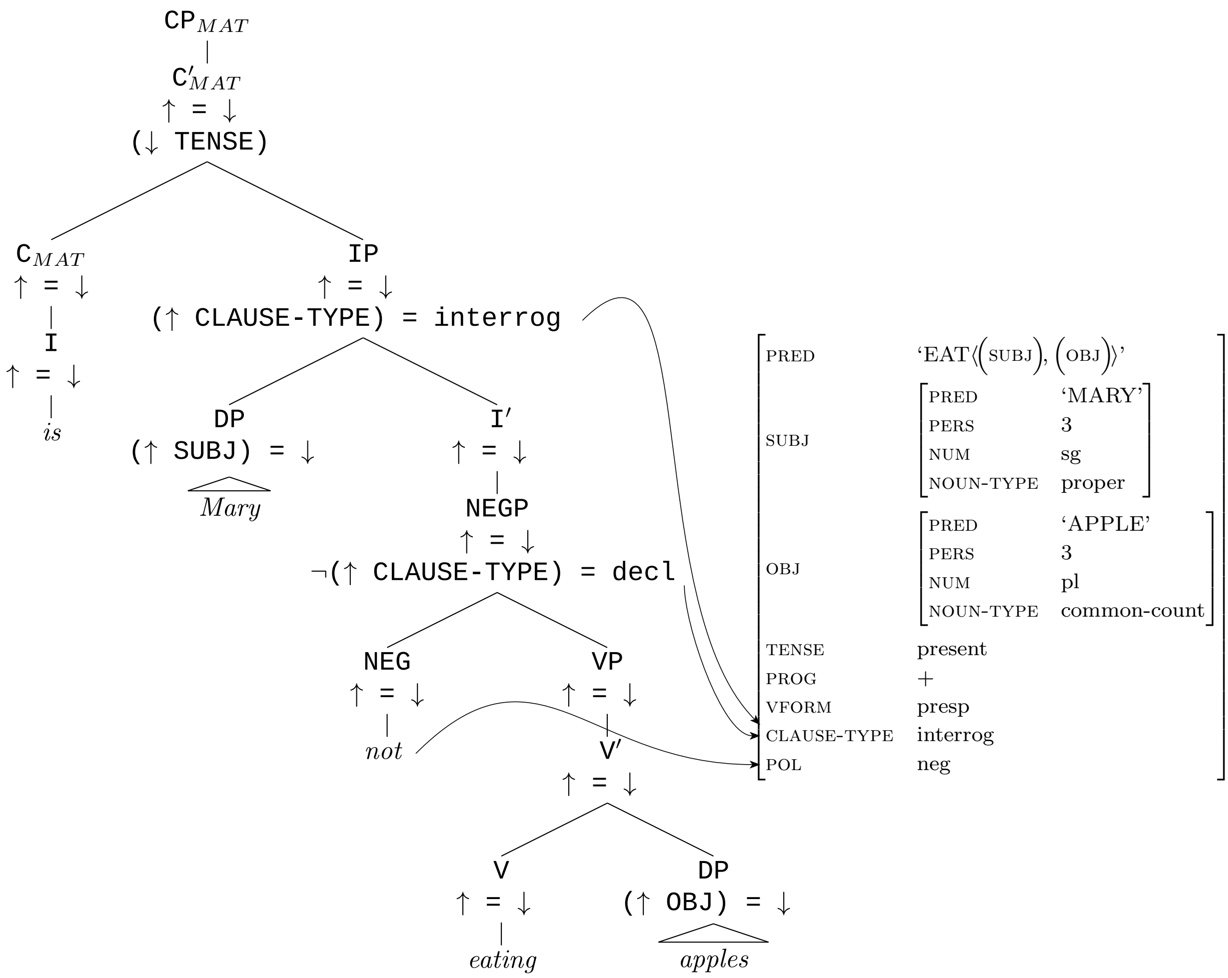
It is somewhat more difficult to handle subject-auxiliary inversion with independent negation following a subject. Such low negation requires an adjustment for the re-write options of I′ in (35a). Instead of substituting I′ with I and NEGP, it must be replaced by NEGP only. The rule must be linked to a clause type feature. The presence of independent negation that is not immediately adjacent to an auxiliary signals a reading of the clausal force as non-declarative, (56).

(56) I′ → NEGP
↑ = ↓
¬(↑ CLAUSE-TYPE) = decl

The rule for NEGP in (35b) itself does not need to be extended. It is noteworthy that sentences with high negation contracted to an auxiliary and low independent negation will result in synonymous feature representations. This is a reasonable approximation within the scope of this chapter. It would be possible, of course, to refine the rules in (54) and (56) to engineer a principled feature difference between the two types of negation.

An example of a sentence with low negation that the model can now parse is *Is Mary not eating apples?*, shown in (57).

(57) Is Mary not eating apples?




The parse of this sentence involves the rule in (56), which re-writes `I′` directly to `NEGP` and expects for the clause type feature some value other than declarative. This option is applicable in the current example because the clause type feature is in fact valued properly as interrogative by the subject-auxiliary inversion rule in (42). The head of the `NEGP`, *not*, values the polarity of the clause as negative. Once again, there is one unique solution for the parse of the string.

Despite the alternative `I′`-rule for low negation, the model still correctly rejects clauses with a finite main verb and direct negation, such as **Mary not eats apples*, (39). The general reasoning regarding this example remains correct. There is no rule available that directly combines a subject with `NEGP` in the relevant, declarative context. However, the model can now, in theory, assign a syntactic structure to the sentence, but not without creating simultaneously a mismatch in clause type features. The absence of subject-auxiliary inversion signals a declarative reading by rule (8), but low negation without an auxiliary precludes this clause type by rule (56).

Finally, I will briefly discuss the phenomenon of negative inversion in Modern English. Subject-auxiliary inversion is obligatory after clause-initial complex negative operators, such as *never*, (58), or `DP`s quantified by *no*, (59).

(58) a. Never did Bill set out on his own to explore the neighborhood.
b. * Never Bill set out on his own to explore the neighborhood.

(59) a. No food could Bill eat without Mary asking for a bite.
b. * No food Bill could eat without Mary asking for a bite.

In broad outline, negative inversion can be analyzed as follows. One can assume that complex negative operators decompose into a negation feature, `POL=neg`, and some other, simpler operator. For instance, the item *never* can be paraphrased as 'not ever', *nowhere* as 'not somewhere', *no* as 'not some,' etc. The isolated negation feature can then be targeted by a syntactic rule that licenses negative constituents in the specifier of $\texttt{CP}_{MAT}$. As before, a clause type feature can be used to regulate the applicability of subject-auxiliary inversion. For this purpose, I introduce a new value for the clause type attribute, called 'non-canonical declarative clauses,' `CLAUSE-TYPE=non-can-decl`. Non-canonical declarative clauses are defined (somewhat tautologically) as declarative matrix clauses that also involve subject-auxiliary inversion. If the rule licensing a clause-initial negative constituent is applied, it will simultaneously introduce such a non-canonical declarative clause-type feature. This is sketched in (60).

(60) $$\texttt{CP}_{MAT} \rightarrow \quad \begin{array}{c} \texttt{XP} \\ \texttt{...} \\ (\downarrow \texttt{POL}) =_{\texttt{c}} \texttt{neg} \\ (\uparrow \texttt{CLAUSE-TYPE}) = \texttt{non-can-decl} \\ \texttt{...} \end{array} \quad \begin{array}{c} \texttt{C'}_{MAT} \\ \uparrow = \downarrow \end{array}$$

The non-canonical clause type feature will be checked in the `IP`-layer. If the respective feature is encountered, the rule introduces a `C`-head dominating an element of category `I`, i.e., an auxiliary. The relevant substitution rule for $\texttt{C'}_{MAT}$ is presented in (61).

(61) $$\texttt{C'}_{MAT} \rightarrow \quad \begin{array}{c} \texttt{C}_{MAT} \\ \uparrow = \downarrow \end{array} \quad \begin{array}{c} \texttt{IP} \\ \uparrow = \downarrow \\ (\uparrow \texttt{CLAUSE-TYPE}) =_{\texttt{c}} \texttt{non-can-decl} \end{array}$$

Thus, the presence of a negation feature in the function corresponding to a phrase in the specifier of `CP` will regulate, via the introduction and constraining of a non-canonical clause-type feature, the mandatory nature of subject-auxiliary inversion in that context.

I added a more comprehensive version of this general sketch of negative inversion to my grammar model. It involves the mapping from rigidly defined lexical items to syntactic structures as well as the semantic computation of the scope of negation. The resulting model constitutes, as far as I know, the first explicit implementation of standardly assumed aspects of negative inversion in LFG. However, the details of my analysis are quite complicated and only of minor importance for the purpose of this chapter. I will therefore not discuss the exact workings of my model for this topic here. The interested reader is referred to Appendix A.

My model of Modern English clause structure can now analyze subject-auxiliary inversion in positive and negative, polar and *wh*-questions as well as after negative initial constituents. The general analyses proposed are quite uncontroversial and could probably be translated straightforwardly into most other syntactic frameworks. In the next section, the model will be extended to cover facts regarding the placement of certain adjuncts.

#### 2.2.1.5 Adverbs and Floating Quantifiers

The position of adverbs differs substantially between sentences with and without auxiliaries. Adverbs are placed in front of finite transitive verbs, (62a), and cannot occur post-verbally before a verbal complement, (62b). However, the adverb diagnostic for the distinction between auxiliaries and lexical verbs is not always as stringent as the negation or inversion contexts. While adverbs typically occur after auxiliaries, (62c), they are often also acceptable in pre-auxiliary position, (62d). The former option, however, is considerably more common (Huddleston and Pullum 2002: 780).

(62)
a. John **probably** visited Spain.
b. * John visited **probably** Spain.
c. John <u>has</u> **probably** visited Spain.
d. $^{ok}$ John **probably** <u>has</u> visited Spain.

There is a vast body of literature on adjunction in general and adverbs in particular. It is impossible to distill from the myriad of different (and often incompatible) analyses one that could be incorporated into my model as a version of the current consensus. Instead, I shall try to sketch, in as general terms as I can, the most important issues in the syntax of adverbs, and justify my decisions for how to address each of them.

First of all, the category "adverb" comprises a large set of distributionally, semantically and formally heterogeneous elements. It is therefore necessary to delineate more carefully what kinds of adverbs should be considered as diagnostics in the context illustrated in (62). I will list the most important properties of the adverbs I am interested in in the following paragraphs.

(1) The relevant adverbs must be able to occur within the vicinity of an auxiliary: The first criterion concerns the syntactic distribution of adverbs, and so it makes sense to describe it within distributional analyses of adverb placement. The earliest influential distributional account of adverbs was presented by Jackendoff (1972). He classifies adverbs according to their possible placement in clause-initial position and before or after the verb phrase of a finite main verb. Hence there are three different adverb positions in his system. This classification was later revised, most notably by Travis (1988) (see also Rochette 1990), to distinguish explicitly between the position before and after an auxiliary, increasing the total number of adverb slots to four. Some distributional analyses also consider adverb occurrences relative to multiple auxiliaries or the influence of an auxiliary's specific morpho-syntactic function. The examples in (63) illustrate a typical instance of a distributional analysis in Jackendoff's and Travis' systems with the adverb *quickly*, which is said to be possible in all positions.

(63) a. (**Quickly**,) John (**quickly**) dropped his coffee (**quickly**). (from: Jackendoff 1972: 49)
b. (**Quickly**,) John (**quickly**) <u>had</u> (**quickly**) finished his meal (**quickly**). (from: Travis 1988: 290)
(Most of my native speaker informants actually disallow the pre-auxiliary position of *quickly*.)

I will not be concerned with clause-initial and clause-final adverbs in any detail here (the first and last slot in Jackendoff's or Travis' systems). Clause-initial adverbs seem to me particularly difficult to study since they show complicating interactions with other factors such as information-structure, intonation and intervening elements. For example, the adverb *loudly* is supposed to be ungrammatical clause-initially (64), but when another element intervenes between this adverb and the subject, the sentence improves considerably, (65). My native speaker informants more or less unanimously confirm this effect.

(64) * **Loudly**(,) he sang the aria. (from: Haumann 2007: 147, ex. (101), *sung* changed to *sang*)

(65) **Loudly**, above everyone else, they recited the sanctification of God's name.
(from: Goluboff, Sascha (2002) *Jewish Russians: Upheavals in a Moscow Synagogue*, p. 53)

Regarding clause-final adverbs, one might perhaps at least mention that some items are banned from this place (e.g., *never*), others can occur in this position but only with a comma intonation (e.g., *evidently*), while still others are perfectly fine there (e.g., *frequently*). Instead of considering these positions further, I merely impose the requirement that the adverb may not be ungrammatical in both positions relative to the auxiliary. This test rules out, for instance, certain adverbs that only encode manner semantics (sometimes called "pure" manner adverbs, e.g., Frey & Pittner 1999), (66).

(66) a. * This running back <u>will</u> {**terribly** / **well** / **fairly** / **elegantly** } play in the NFL.
b. * This running back {**terribly** / **well** / **fairly** / **elegantly**} <u>will</u> play in the NFL.

Some other, non-manner adverbs are excluded by this criterion as well.

(67) a. * The disease <u>will</u> {**early** / **more** } affect him.
b. * The disease {**early** / **more** } <u>will</u> affect him.

Adverbs like these must be ignored because their inability to appear anywhere in front of the full lexical verb deprives them of any diagnostic value for the indication of auxiliary or main verb positions.

(2) The relevant adverb must not be able to occur between a transitive main verb and its complement: The second criterion for the identification of appropriate adverbs is another distributional test. The adverb must not be natural between a lexical head like a (transitive) main verb and its complement. The overwhelming majority of adverbs are in fact unacceptable in this position. This may explain why this slot has not been considered in great detail in distributional analyses (but see Haumann 2007: 30-32), and it has not sparked a distributional classification of adverbs. The test works best with phonologically light complements to avoid heavy NP shift, such as the bare plural *shoes* in (68).

(68) a. John will buy ______ shoes.
b. $^{ok}$ John will buy {**even** / **just** / **merely**} shoes.
c. * John will buy {**probably** / **already** / **certainly**} shoes.

This test is mainly designed to exclude adverbs that have a high probability of combining with DPs to form complex nominal constituents such as *even*, *just* or *merely* above. This is necessary because the ambiguity between nominal and clausal modification can considerably alter the statistical information on the positional difference between auxiliaries and full lexical verbs conveyed by these adverbs in natural corpus data.

(3) The relevant adverb must not allow for a manner reading out of context: I decided to discard all manner adverbs from my study. By "manner adverb" I mean any adverb that can in principle function as an answer to a 'in what way' question. The reasons for the categorical exclusion of manner adverbs are the following. Some manner adverbs are acceptable after but not in front of an auxiliary, (e.g., Wyner 1998), (69).

(69) a. This running back <u>will</u> {**passionately** / **brilliantly** / **skillfully**} play in the NFL.
b. * This running back {**passionately** / **brilliantly** / **skillfully**} <u>will</u> play in the NFL.

Such adverbs would be ideal diagnostics to distinguish between auxiliaries and main verbs distributionally. However, their manner semantics is almost always incompatible with the stative sense of possessive *have*. They can thus not be considered here, (70).

(70) * John {**carefully** / **thoroughly** / **expertly**} has a new idea.

Other manner adverbs are ambiguous between their strict manner reading and some kind of "oriented" reading. The fact that all of these ambiguities involve an adverb with a potential manner interpretation may suggest that the oriented reading somehow derives from the manner reading as its basis (e.g., Kubota 2015). Ambiguities of this nature, first identified in Jackendoff (1972), have frequently been discussed, for instance as a contrast between manner and subject-oriented (e.g., McConnell-Ginet 1982) or, closely-related, agent-oriented readings (e.g., Ernst 2002: chapter 2), manner and speaker-oriented (or "pragmatic") readings (e.g., Bellert 1977), and several others. The following example illustrates the ambiguity between manner and subject orientation with the adverb *rudely*.

(71) John **rudely** asked a question.
a. manner: John asked a question in a rude way.
b. subject-oriented: It was rude of John to ask a question.

The semantic difference is truth-conditionally relevant. John could have asked a question in a very polite manner but in a setting where no question should be asked. In this case, sentence (71) would be false under the manner reading (71a) but true under the subject-oriented reading (71b). Crucially, the interpretation of such adverbs is subject to a word order effect. Specifically, adverbs of this kind in front of auxiliaries allow for a non-manner reading only. This is illustrated for subject-oriented adverbs in (72) and for speaker-oriented adverbs in (73).

(72) a. John **cleverly** <u>had</u> talked to the police.
i. *John had talked to the police in a clever way (manner).
ii. John was clever to talk to the police (subject-oriented).
b. John <u>had</u> **cleverly** talked to the police.
i. John had talked to the police in a clever way (manner).
ii. John was clever to talk to the police (subject-oriented).

(73) a. John, **frankly**, <u>had</u> talked to the police.
i. *John had talked to the police in a frank way (manner).
ii. The speaker is frank in asserting that John had talked to the police (speaker-oriented).
b. John <u>had</u> (,)**frankly**(,) talked to the police.
i. John had talked to the police in a frank way (manner).
ii. ?? The speaker is frank in asserting that John had talked to the police (speaker-oriented).

As these examples show, ambiguous manner adverbs are not actually in free variation in pre- and post auxiliary position. Rather, meaning differences may explain a speaker's choice to put the adverbs in a specific place. Since it is not easily possible to tell which meaning was intended in historical texts, adverbs of this type must be excluded from the present study as well. In short, manner adverbs are categorically unsuitable for the goals of this chapter.

(4) The relevant adverb must modify propositions: The adverbs I am interested in should, at least *prima facie*, allow for an analysis as propositional modifiers. They combine with propositions to yield propositions, i.e., they are elements of type $< t, t >$ (Halverson 1983). I will simply represent the meaning of such adverbs as

ADV($p$), where $p$ stands for the proposition modified by the adverb. The specific semantics of the adverb will not be of any concern here. They would have to be formalized further in a logical language. For instance, *always* could be formalized in a temporal logic such that the proposition it modifies holds at all times; *necessarily* could be analyzed in an intensional system such that the proposition it modifies holds in all possible worlds, etc. For an example of the way I represent the meaning of a relevant adverb, see the semantic contribution of the adverb *never* in (5) in Appendix A.

It is sometimes doubted that any specific class of sentential adverbs could uniformly consist of propositional modifiers (e.g., Delfitto 2006: 90). In fact, it is usually difficult to determine the semantic type of any adverb. One relatively foolproof argument for an adverb's status as a propositional modifier consists in demonstrating that it participates in scopal ambiguities (for instance with generalized quantifiers in argument positions, e.g., Ruys and Winter 2011: 168). This is illustrated in (74).

(74) a. *The school festival used to involve a clown that all children found funny.*
All students **sometimes** laughed.
$sometimes(\forall x[student(x) \rightarrow laugh(x)])$

b. *Even the best students couldn't always pull themselves together.*
All students **sometimes** laughed.
$\forall x[student(x) \rightarrow sometimes(laugh(x))]$

In (74a), the subject scopes below the adverb yielding a cumulative reading ('it is sometimes the case that all students ...') whereas (74b) has the subject with wide scope resulting in a distributive reading ('for all students, sometimes ...'). Hence, an entire proposition must apply to the adverb, as sketched in the semantic formulas in the above examples. In contrast, manner adverbs do not engage in such scope ambiguities.

(75) *The teacher used to make a lot of bad jokes.*
All students **politely** laughed. (no scope ambiguity)

Unfortunately, it is difficult to use scope ambiguities as a reliable test for an adverb's semantic type. The scope ambiguities are often so subtle that it becomes difficult to see truth-conditionally relevant meaning differences. Moreover, the test requires the adverb to quantify over a (set of) elements like times, worlds etc. But it is conceivable that an adverb is not quantificational in that sense and still scopes over a proposition. Hence the test may identify some but not all adverbs of type $< t, t >$. A different criterion has been suggested to distinguish propositional modification from other modifiers, namely the distinction between *de re* and *de dicto* readings in opaque contexts (Heny 1973). But judgments for this test are frequently even more subtle than those for scope ambiguities with quantifiers. I shall therefore not pursue a discussion of this diagnostic any further here.

Instead, I will use a simple paraphrase test as a proxy for the desired status of the adverb as a propositional modifier. Specifically, the relevant adverbs must be able to occur in the paraphrases 'it is $ADV$ the case that...' or 'it is $ADV$ true that... .' These are relatively natural paraphrases for a wide scope reading of an adverb over the proposition introduced by the complementizer. Some adverbs that meet this criterion are shown in (76a). These are thus the adverbs that I will consider here. Manner adverbs, on the other hand, are generally impossible in this construction, (76b). However, if the adverbs are ambiguous between a manner and some oriented, non-manner reading, they may occur in these paraphrases with the non-manner reading prevailing. Nevertheless, like all other manner adverbs, such elements will be excluded here for the reasons explained for the third criterion above, (76c). Finally, there are several non-manner adverbs that cannot occur in the paraphrases. These are the adverbs that are effectively eliminated by this test. Their semantics are presumably fundamentally different from those of the other classes, (76d).

(76) a. It is { **necessarily** / **really** / **perhaps** / **fortunately** } the case / true that ...

b. *It is { **gently** / **nervously** / **terribly** / **slowly** } the case / true that ...

c. It is { **clearly** / **oddly** / **obviously** / **seriously** } the case that / true that ...
* manner: in a clear / odd / obvious / serious manner
$^{ok}$ non-manner

d. *It is { **here** / **today** / **botanically** / **annually** } the case / true that ...

To summarize, the adverbs that remain as potential candidates for the corpus study in the second half of this chapter are non-manner adverbs that can occur immediately before and / or after an auxiliary, but not normally in between a full lexical verb and its complement, and in a *be the case that...* paraphrase. Some examples of such adverbs are shown in the example below.

(77) a. Mexicans {**frequently** / **never** / **certainly**} will answer polls the way they think sounds best.
b. ≈ Mexicans will {**frequently** / **never** / **certainly**} answer polls the way they think sounds best.

Note that there are still some uncontrolled factors that limit the syntactic occurrence even of such items. Pre-auxiliary adverbs are said to be preferred with stress and / or focus on the auxiliary, (78). Further, pre-auxiliary adverbs can often take only wide scope over a modal auxiliary whereas post-auxiliary adverbs remain ambiguous, (79). Similar complications regarding scope arise with negation.

(78) a. John is **often** [in his OFfice]$_{[+\text{foc}]}$
b. John **often** [IS]$_{[+\text{foc}]}$ in his office.
(from: Wilder 1997*a*: 327-8)

(79) a. John **frequently** could lift 200 pounds.
i. ‘John was frequently able to lift 200 pounds.’
ii. # ‘John was able to lift 200 pounds several times (in a row)’
b. John could **frequently** lift 200 pounds.
i. ‘John was frequently able to lift 200 pounds.’
ii. ‘John was able to lift 200 pounds several times (in a row)’
(from: Engels 2012: 4, ex. (1.6.))

However, phonological or information-structural information is extremely difficult to recover from written texts, and scope ambiguities with modals or negation seem to occur so infrequently in corpus data as to be negligible. I will therefore not restrict permissible adverbs any further but accept phonology, information structure and scope ambiguities as possible sources of noise in my data.

Having identified the relevant adverbs for this study, I will now consider the best way to analyze them syntactically. In principle, there are two ways to model the initial observation that adverbs and verbs can be variably positioned relative to each other. The first possibility is that adverbs are fixed in one position (given an information-structural and syntactic context), and verbs are placed in a number of different positions before and after the adverbs. One of the most influential accounts of this option is the cartographic functional sequence of adverbs (Cinque 1999). Couched in derivational terms, it says (i) that verbs can move to higher or lower head positions of projections whose specifiers can be realized as adverb phrases, (ii) that these projections are rigidly ordered in the syntax as a hierarchy, and (iii) that this hierarchy is universal across all languages. A typical example of data that is cited in support of these claims is shown in (80).

(80) *Italian*

| | | | | | | |
|---|---|---|---|---|---|---|
| a. | Non hanno | mangiato | **mica** | | **più** | |
| b. | Non hanno | | **mica** | mangiato | **più** | |
| c. | Non hanno | | **mica** | | **più** | mangiato |
| | not they-have | (eaten) | not | (eaten) | any-longer | (eaten) |

‘They didn’t eat any longer’ (adapted from: Cinque 1999: 47, examples (7)-(11))

In this example, the adverb *mica* must necessarily precede the adverb *più*; the inverse ordering is ungrammatical in any configuration. The non-finite main verb can come before, in between or after this adverb sequence. This may indicate that “verbs, not AdvPs, can occupy different positions within a certain [...] ‘space’” (ibid.: 51). Adopting this analysis for adverb placement in the English examples (62c) and (62d), one could thus claim that there are two positions for the finite perfect auxiliary *has* available, one before and one after the invariantly positioned adverb *probably* (ibid.: 109).

The second option is to assume that verbs are fixed in one position (given an information-structural and syntactic context), and that adverbs are placed in a number of different positions around them. There is currently no one model of this approach to adverbs that could be regarded as prototypical and dominant in the field. Instead, a large number of different proposals exist that roughly agree on variable adverb positions but differ substantially in detail (e.g., Jackendoff 1972, Potsdam 1998, Ernst 2002, Ramchand and Svenonius 2014; for an overview, see Haumann 2007: 71-102). In my understanding, these models all share the following basic assumptions. (i) Adverbs come in a (relatively) small number of classes that are ultimately determined by their lexical semantics, and specifically by the semantic type they select for. For instance, some adverbs modify propositions and are sentential in that sense, so called S-adverbs, while other adverbs are related to events or processes expressed by the verb phrase, so-called VP-adverbs (terminology from Jackendoff 1972 and

Potsdam 1998, similarly Dalrymple 2001: 269-74, who discusses the difference between sentential *obviously* and manner *skillfully*; Ernst 2002 and Ramchand and Svenonius 2014 have a considerably more complex ontology). (ii) Adverbs are not sequentially ordered in a hierarchy. Rather, particular orders follow from a number of unrelated factors. Perhaps most importantly, rules of semantic composition (e.g., Ernst's (2002) Fact-Event Object Calculus, Ramchand and Svenonius' (2014) "sortal domains") force a particular syntactic position for an adverb of a particular type to combine with an appropriate semantic argument. Such compositional rules are used, for instance, to account for the general order of S-adverbs before VP-adverbs, as in (81) (see also Ernst 2002: 127 for examples of the same kind within his system; see Bulkley Cobb 2006 for a feature-based implementation in LFG).

(81) a. Hulk Hogan $[\text{evidently}]_{\text{S}}$ $[\text{completely}]_{\text{VP}}$ annihilated his opponent.
b. * Hulk Hogan $[\text{completely}]_{\text{VP}}$ $[\text{evidently}]_{\text{S}}$ annihilated his opponent.
(from: Potsdam 1998: 402, ex. (18))

Particular adverb orders can also be determined by scope. This is illustrated for the adverbs *intentionally* and *never* in (82) below.

(82) a. The speaker **intentionally never** strays from the topic.
b. The speaker **never intentionally** strays from the topic.
(from: Ernst 2002: 131, ex. (3.127))

Sentence (82a) is false if the speaker strays from the topic, whereas sentence (82b) may be true in this case as long as he or she does not do so intentionally. The position of the two adverbs relative to each other thus leads to different scopal interpretations. Proponents of variable adverb positions also maintain that not all adverbs actually show ordering effects. Instead, there are cases of free adverb ordering, for which it is impossible or at least extremely difficult to find contexts that would induce a truth-conditionally relevant meaning difference. An example is shown below.

(83) The rebels have $\left\{\begin{array}{l}\text{now perhaps}\\ \text{perhaps now}\end{array}\right\}$ surrendered. (from: Potsdam 1998: 400, ex. (13a))

(iii) Finally, models of variable adverb positions tend to employ as their formal mechanism for adverb placement Chomsky-adjunction to specific syntactic categories. If one were to adopt an analysis along these lines, the English examples (62c) and (62d) would involve two adverb positions, one before and one after the invariantly positioned auxiliary *has*.

Obviously, the crucial question one should ask to decide between these two alternatives is whether the evidence for adverb default orders, and hence for a hierarchy of syntactic projections in models with fixed adverb and variable verb positions, outweighs the evidence for flexible adverb orders, and thus at least partially non-syntactic mechanisms of adverb placement such as scope resolution in models where the adverb is flexible and the verb is fixed. In order to base my answer to this question at least on some empirical material, I decided to conduct a grammaticality judgment experiment.

The experiment[3] made use of the three non-manner adverbs *possibly*, *often* and *no longer*. These adverbs are of the type identified above as relevant for the current study. Furthermore, Cinque provides an explicit ordering relation between them by identifying dominance of the possibility modality (*possibly*) over the frequentative(I) aspectual (*often*) over the terminative aspectual (*no longer*) heads, as shown below.

(84) [*possibly* $\text{Mod}_{\text{possibility}}$ [ ... [*often* $\text{Asp}_{\text{frequentative(I)}}$ [ ... [*no longer* $\text{Asp}_{\text{terminative}}$ ... ]]]]
(adapted from: Cinque 1999: 106, ex. (92))

Note that Cinque includes a second frequentative head lower in the structure, called aspect frequentative(II). Its specifier can also be realized by the adverb *often*. The adverb *often* of the higher head takes wide scope over the whole proposition, can be paraphrased as "it is often the case that ..." as well as by binding a bare subject DP with 'most' (*Texans often drink beer* ≈ 'Most Texans drink beer,' ibid.: 26). This is the kind of *often* that I am interested in here. In contrast, the adverb *often* of the lower head "indicates the *repetition* of the act denoted by the verb" (ibid.), i.e., it scopes over the event expressed by the verb phrase only, and usually occurs post-verbally (*Texans drink beer often* ≈ 'Typically all Texans drink beer more times than is usually the norm,' ibid.: 26-7). This is not the kind of interpretation of *often* targeted in my experiment. It is therefore fair to

[3]The raw data in the form of 100 survey responses and a summary excel sheet can be found on the Dissertation DVD under `Chapter 2 - Possessive Have/01 Adverb Scope Study`.

say that Cinque would predict the default word order *possibly* - *often* - *no longer* in sentences in which all three adverbs appear. Models of variable adverb placement, on the other hand, would predict that the relative order of these three adverbs is not inherently fixed but depends on their relative scope.

The material for the experiment consisted of six test sentences and associated contexts. The test sentences were all passives involving the auxiliary *are* followed by the three adverbs in a row and a past participle. The contexts were meant to induce a particular scopal reading for the three adverbs, which might be mirrored by a corresponding surface word order. The first context targeted the scope corresponding to the word order predicted by Cinque, *possibly* - *often* - *no longer*. The other five contexts were designed to work best with one of the remaining scopal orders. Example (85) shows the six contexts and test sentences with the targeted adverb surface orders reflecting the most appropriate underlying scope.

(85) a. Context: People who are looking to purchase a house absolutely always had to pay a 20% down-payment in the past. However, economists believe that in many situations banks do not require any down-payments anymore at all! Surveys among banks are now being conducted to find out if that's really true.
Down-payments are **possibly often no longer** required on conventional mortgages.

b. Context: Planets outside of our solar system, so-called 'exoplanets,' have never been photographed before. Scientists will now attempt to take pictures of them for the first time. Their photography experiment will begin in December, and then continue once a week for the next 5 years. Hence, there will be very many situations in which the scientists might perhaps be able to see exoplanets directly.
Exoplanets are **often possibly no longer** blocked from our direct view.

c. Context: Newspapers have reported, incessantly and all the time, that cigarettes were the number one cause of cancer. But recently the incidence of lung cancer in smokers has declined sharply, perhaps as a consequence of the introduction of less harmful electronic cigarettes. Research is now under way to test if this is also reflected in a lower number of newspaper articles that are critical of smoking.
Cigarettes are **possibly no longer often** mentioned as a dangerous drug in newspapers.

d. Context: Based on certain evidence, one was justified in assuming for a long time that vaccines played a role in the development of autism. However, most (though not all) vaccines have now been proven beyond any reasonable doubt to be absolutely safe. In a great number of cases, it really is not plausible anymore to believe that vaccines contribute to neurodevelopmental disorders in any way.
Vaccinations are **often no longer possibly** associated with autism.

e. Context: Nurses used to face the possibility of catching a dangerous disease from their patients in many different situations: They could potentially get infected, for example, when they performed physical examinations, when they took someone's temperature, when they repositioned a patient, and so on. However, a new drug for nurses has been developed. It will definitely prevent 100% of all care-associated infections from now on.
Nurses are **no longer often possibly** exposed to communicable diseases.

f. Context: The international catalog of diseases was very imprecise. It might have led to diagnosis mistakes. Some doctors say that 3% of all diagnoses were wrong, but others believe mix-ups might have occurred as frequently as 8 out of 9 times. Nobody knows how often wrong diagnoses were made, but it is clear that it could have been frequently. Luckily, 100% of all confusion has now been removed from the disease catalog. Diseases definitely won't be misdiagnosed anymore.
Diseases are **no longer possibly often** diagnosed incorrectly.

The experiment took place in the form of an online survey designed with the Qualtrics platform. 100 participants were recruited from the online crowdsourcing marketplace Amazon Mechanical Turk. They were asked to take the survey only if they were native speakers of American English and their IP address was tracked to United States territory. Every participant was paid 75 US cents for their participation. On average, the experiment lasted about 6 minutes.

The experiment proceeded as follows. Participants were told that they were taking part in a science comprehension study. They first had to order adverbs in two test questions where the relative adverb order was uncontroversially fixed (speech act adverb before modal adverb, similar to (81) above; *not* preceding negative polarity item *any longer*). They were then randomly presented with two out of the six contexts in (85). They had to answer a very simple comprehension question about the context to ensure that they had read it attentively. Next, the participants were asked to select one of the six possible relative orders of *possibly*, *often* and *no longer* in the passive test sentences and confirm their answer by selecting the same word order again. The six possible answers were presented in a randomized order in both trials. If a participant either did not answer the comprehension question correctly or gave two different adverb word orders, their response was not considered for evaluation for that test item. Finally, participants were asked to rate the grammaticality of the sentence they had formed on a Likert scale ranging from 0 to 6. A filler was included between the two test items.

Informally speaking, I designed the experiment to decide between one of the following two hypotheses. If the order between adverbs is fixed, which could be attributed to a syntactic hierarchy of projections, then one should expect one word order (presumably Cinque's *possibly* - *often* - *no longer*) to emerge as dominant across the six test conditions. If, on the other hand, the order between adverbs is not inherently fixed, but rather due to surface reflections of underlying semantic scope, then one would predict a different, namely the targeted, word order to be predominant for each of the six contexts.

The results of the experiment are as follows. Firstly, a surprisingly large number of responses had to be rejected. In total, I collected 200 responses (2 randomly chosen test items for each of the 100 participants). Of those, only 106 could be accepted. The remainder of the responses involved incorrect answers to the comprehension question or the submission of two contradictory word orders for the same test sentence. This may indicate either that the task was quite difficult or else that the recruited participants did not take the task seriously but just answered the questions quickly to receive their compensation. As a result, I skewed the probability with which certain contexts were displayed to the participants about halfway through the experiment to collect a comparable number of acceptable responses for each of the six contexts.

Secondly, the 106 accepted responses showed that the relative order of the three adverbs did indeed correlate with the scopal readings implied by the contexts. For each of the six context in (85a) - (85f), the word order targeted for that context emerged as the most frequent. The only exception was context six, (85f), for which the number of responses of the targeted word order shared the highest rank with two other word orders. These results are illustrated by the barplots in figure 2.1.

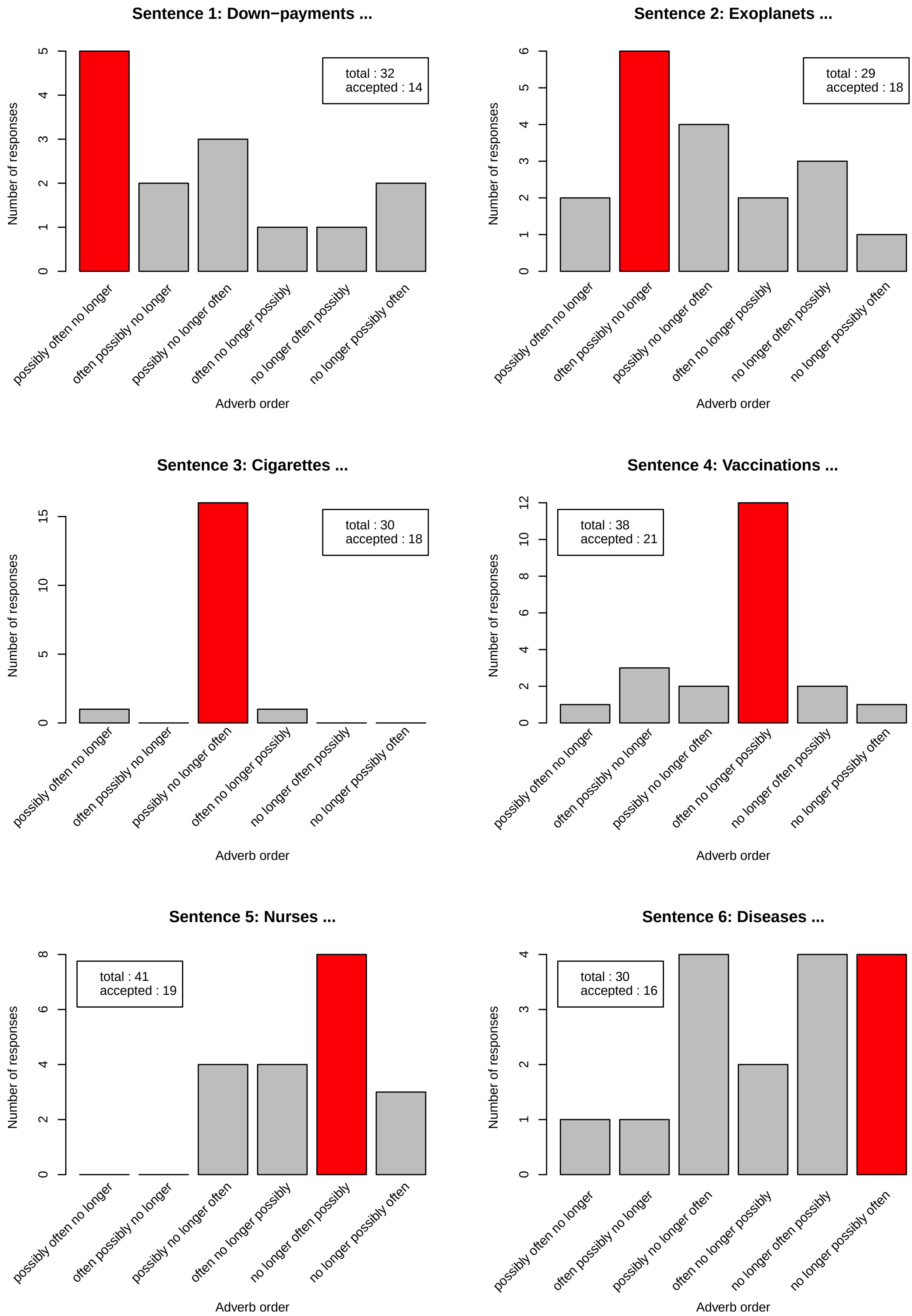



Figure 2.1: Results of the adverb ordering task

The six barplots correspond to the six test sentences in (85a) - (85f). Each plot indicates the six possible orderings of the three adverbs *possibly*, *often* and *no longer* on the x-axis and the frequency with which participants chose a particular word order on the y-axis. The red bar highlights the word order targeted for that context. For each sentence, the red bar has the largest number of responses. It is possible to test statistically if the targeted word order appeared significantly more often than would be expected by random chance. Assuming a binomial distribution in which each of the six word orders has a probability of 1/6 and the number of accepted responses corresponds to the number of trials (i.e., the participants were randomly picking a word order), one can calculate the probability of finding the observed number of responses or more as 1− the cumulative probability function up to the observed value minus 1 (one sided test). For example, the reasoning for context 1 is as follows: If the probability of choosing the order *possibly* - *often* - *no longer* is just 1/6, then the probability of finding 5 or more such responses out of a total of 14 accepted responses would be $p = 0.069$. Similarly, the probability of finding at least the observed number of responses of the targeted word order would be $p=0.065$ in context 2, $p< 0.001$ in context 3, $p< 0.001$ in context 4 and $p=0.008$ in context 5. Solely context 6 has a substantially higher p-value at $p=0.271$. Sentence 6 might be associated with a target word order that is genuinely harder to intuit than the others, or its context might have been poorly designed so that it did not provide enough information on the intended scope. Even though not all targeted word orders reached a significant value at the 5% level, and context 6 did in fact display quite a high chance of random guessing, the overall low p-values and consistent biases towards the targeted word orders suggest that participants were not picking word orders at random but were instead guided by the information provided in the contexts. This, then, lends support to the presence of a scope effect in the determination of surface adverb order and thus to the hypothesis that the order between adverbs is not inherently fixed. Furthermore, it is clear that the frequencies of the adverb word orders is very different across the six test items. For example, a Fisher's Exact test yields a significant difference between the frequencies of Cinque's default order *possibly* - *often* - *no longer* and all other word orders combined across the six contexts, $p=0.021$. Therefore, it seems very unlikely that the six response patterns are drawn from the same distribution. This argues at least somewhat against the hypothesis that one word order could constitute a default.

The third main finding of my experiment is that there may nevertheless be some evidence for a default order in the data. However, if there is a default order at all, it must be *possibly* - *no longer* - *often* rather than Cinque's *possibly* - *often* - *no longer*. This sequence is the most frequent order overall (33 of 106 responses) and significantly more common than would be expected by random chance, $p< 0.001$. The context targeting this word order (sentence 3) yielded the least ambiguous result (16 targeted responses out of 18 responses) and the word order was also the strongest competitor to the targeted order in 4 out of the remaining 5 contexts. However, the principal reason why the word order *possibly* - *no longer* - *often* appears so frequently is precisely because context 3 yielded such unambiguous results. If test sentence 3 is removed from the data, this word order does not occur significantly more frequently than expected by random chance (17 of 88 responses), $p=0.394$. The evidence provided by this experiment in favor of a default adverb order thus remains relatively weak overall.

Fourth, one might speculate that a targeted word order was picked in its respective context simply as the result of some kind of puzzle solving activity but that one particular default order, possibly Cinque's *possibly* - *often* - *no longer*, actually sounded much more natural than all the others. Hence, the default order might be detected not through a task that forces participants to order adverbs in a context but rather through a grammaticality judgment task. To evaluate this conjecture, I calculated from the 106 accepted responses the mean acceptability rating for each of the six adverb orderings provided by the participants. The mean acceptability rating was also calculated for the filler item as a control. For the filler, participants had to order only two adverbs relative to each other, namely *no longer* before *slowly*, as shown in (86).

(86) Context: Due to a software bug, the Mars rover moved at an incredibly slow pace. Scientists at NASA have now fixed the problem. In the future, the rover will move much faster.
The Mars rover will **no longer slowly** move across the Martian surface.

The filler item showed a much higher success rate than the test items. 90 of 100 responses could be accepted; only 10% involved the untargeted word order, an incorrect answer to the comprehension question or the submission of two incompatible orders. The result of the grammaticality judgment task are shown in the barplot in figure 2.2. The six word order variants and the filler are shown on the x-axis, the height of the bar on the y-axis represents the mean of all the judgments given for that variant (with 6 being the most acceptable, natural or grammatical value), and the error bars represent one standard deviation above and below that mean. The ratings of every word order variant were compared to the ratings of every other word order variant to test for statistically significant differences between their means using Wilcoxon rank sum tests. These statistics are presented in table 2.1.

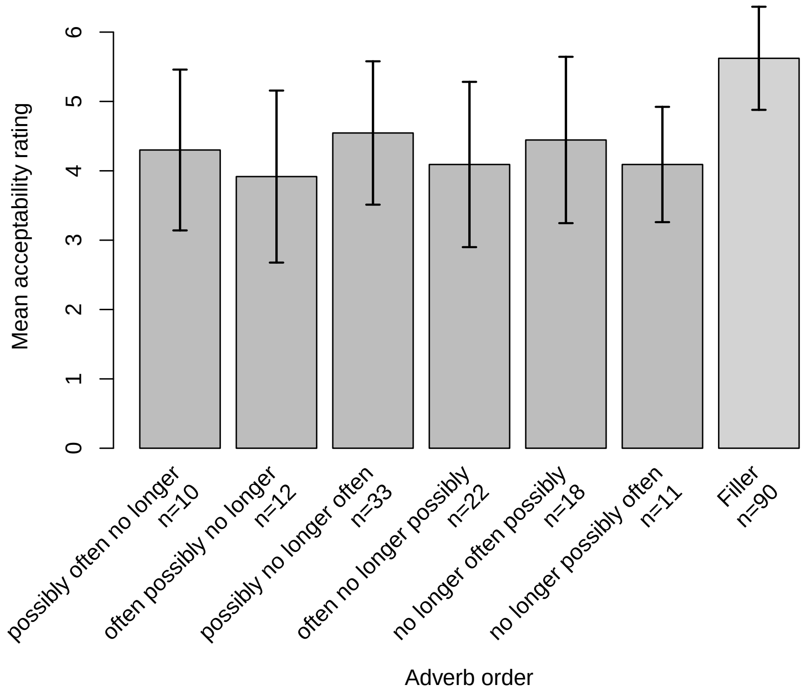



Figure 2.2: Grammaticality judgments for sentences with three adverbs and the filler

| | often possibly no longer | possibly no longer often | often no longer possibly | no longer often possibly | no longer possibly often | Filler |
|---|---|---|---|---|---|---|
| possibly often no longer | W=72.5<br>$p$=0.406 | W=147.5<br>$p$=0.609 | W=124<br>$p$=0.566 | W=84.5<br>$p$=0.804 | W=63.5<br>$p$=0.553 | W=137<br>$p$<0.001*** |
| often possibly no longer | - | W=138<br>$p$=0.111 | W=122.5<br>$p$=0.731 | W=83<br>$p$=0.282 | W=62<br>$p$=0.819 | W=120.5<br>$p$<0.001*** |
| possibly no longer often | - | - | W=444.5<br>$p$=0.147 | W=309.5<br>$p$=0.806 | W=230<br>$p$=0.174 | W=585.5<br>$p$<0.001*** |
| often no longer possibly | - | - | - | W=165.5<br>$p$=0.367 | W=122<br>$p$=0.984 | W=287<br>$p$<0.001*** |
| no longer often possibly | - | - | - | - | W=118.5<br>$p$=0.376 | W=334.5<br>$p$<0.001*** |
| no longer possibly often | - | - | - | - | - | W=87<br>$p$<0.001*** |

Table 2.1: Wilcoxon rank sum tests for each word order pair

The six orderings of three adverbs all cluster around a mean acceptability score of about 4. Thus, participants seem to have found sentences involving three adverbs marginally acceptable irrespective of their specific word order alignment. Indeed, there are no statistically significant differences in the mean acceptability ratings for any pair of these orders. The control sentence with only two adverbs, in contrast, received an average rating of about 5.5. The filler sentence therefore seems to have been assessed as much more acceptable than any of the other sentences. This is also shown by the significance tests, which reveal significant differences in mean acceptability between the filler and each of the six configurations of three adverbs. The significant differences between the filler and the test sentences points towards an effect of the number of adverbs involved in a sentence. Sequences with 3 adverbs appear to be consistently less natural than sentences with a sequence of only 2 adverbs. Conversely, the fact that all of the three adverb orderings are equally acceptable makes it unlikely that any one of these alignments, in particular Cinque's presumed sequence, could constitute a default.

In conclusion, my experiment provides some limited evidence against a hierarchical order of adverb sequences and in favor of scope effects in the determination of adverb orderings at least for the kind of adverbs that I am interested in here. Interestingly, Cinque himself recognizes in his concluding remarks that "many (perhaps most) of the relative orders among functional elements may ultimately reduce to scope relations among what we can take to be different semantic operators" (1999: 141). My experiment suggests that Cinque's reflection on this point is in fact correct. I will therefore proceed to implement adverb placement with a traditional adjunction mechanism rather than with a syntactic hierarchy of projections.

Before I continue to extend my grammar fragment accordingly, I would like to mention a few more arguments against the implementation of a model of adverbs that involves invariant adverb and flexible verb positions. First of all, there appear to be purely empirical problems with Cinque's model. It seems to make numerous wrong predictions, at least for English (e.g., Edelstein 2012: 15-24, Zyman 2012). Next, adjunction has traditionally been, and is still today, the most widely assumed formal device to integrate adverbs into the syntactic structure. It is also almost universally used in the LFG literature (see for example the guidelines for functional annotations to phrase structures rules in Bresnan 2001: 102, examples (21e), (22e)). Since it is my goal to construct as uncontroversial a model of Modern English clause structure as possible, it makes sense to me to follow the majority view regarding this issue. Finally, and perhaps most importantly, it actually turns out to be extremely difficult to implement Cinque's theory in a formally rigid manner. Specifically, verb placement rules under different syntactic heads (i.e., head movement in derivational frameworks) are context-sensitive operations. For example, in the model developed here, the C-head can be re-written as I if the clause type is interrogative, (43), or non-canonical declarative after negative initial constituents, (61). If no contexts were explicitly specified at all, nothing would seem to prevent just about any syntactic head to re-write to the verb (i.e., a verb could move too far or not far enough in derivational frameworks). The question is then which contexts account for the variable distribution of verbs across the twenty or so head positions proposed by Cinque (i.e., what features drive verb movement in derivational frameworks). Such context-sensitive rules impose quite a high computational burden from a grammar-engineering perspective. Thus, if I were to attempt to incorporate a version of Cinque's adverb hierarchy into my grammar model, I would run into problems of mathematical explicitness quite quickly.

At long last, I will now present the implementation of adverbial syntax in my grammar model. I will first point out several limitations of the model developed here. (i) The term "adverb" (hence also the adverb phrase it projects, ADVP) refers only to the rather capacious class of adverbs delineated at the beginning of this section. (ii) There is no attempt to provide fine-grained semantic denotations for these adverbs. I will simply assume that the relevant adverbs can be analyzed, at least superficially, as propositional modifiers. (iii) I will not provide rules of semantic composition even for sentences that contain only instances of these adverbs. (iv) Since adverbs modifying elements other than propositions are not considered here, the model does not implement ordering restrictions between adverbs that emerge as a consequence of their selecting for divergent kinds of semantic arguments either. (v) Similarly, the model cannot account for the semantic relation between manner adverbs and oriented adverbs and their related word order effects. It might be possible to associate certain positions with semantic resources that could derive the ambiguity between different adverb readings, but such an analysis would go far beyond the scope of this chapter. (vi) Since detailed denotations involving times, worlds, situations etc. for the adverbs relevant here are not developed, it is also impossible to model their interaction with modals and negation adequately. Furthermore, modality and negative polarity are simply implemented as features in my model, which would require a substantial revision if their relation to adverbs were to be more thoroughly explored. (vii) The model can handle pre-auxiliary and post-auxiliary adverb phrases but other adverb positions are not included.

I will, however, make an attempt at implementing a rudimentary mechanism for scope relations between the adverbs considered here because scope emerged as the most relevant factor to account for the relative order between these adverbs during the preceding discussion. The scope mechanism will make use of two functional constraints, which will be combined to yield a new set operator, as described in the documentation of the standard LFG grammar engineering platform, Xerox Linguistic Environment (XLE) (Crouch et al. 1993). The first constraint simply asserts that a scope relation exists between two functions, as defined in (87).

(87) *scope relation*
$f$ scopes over $g$ ($f <_s g$) if and only if $<_s$ is a set and $f$ is a member of the set ($g <_s$).

This constraint can be used to place functions in a second function's set $<_s$, which contains elements that scope over it. The second constraint is called "head precedence." It is defined in (88) below.

(88) *head precedence*
$f$ head-precedes $g$ ($f <_h g$) if and only if the node containing $\phi^{-1}$($f$ PRED) precedes the node containing $\phi^{-1}$($g$ PRED).

The inverse relation $\phi^{-1}$ associates the functional structure with the constituent structure. The definition of head precedence thus states that the node introducing the `PRED` feature of a function $f$ precedes the node introducing the `PRED` feature of a function $g$ in the constituent structure. With those two constraints, one can now define a special kind of set membership, which I call scope-relational set membership. The concept is defined in (89).

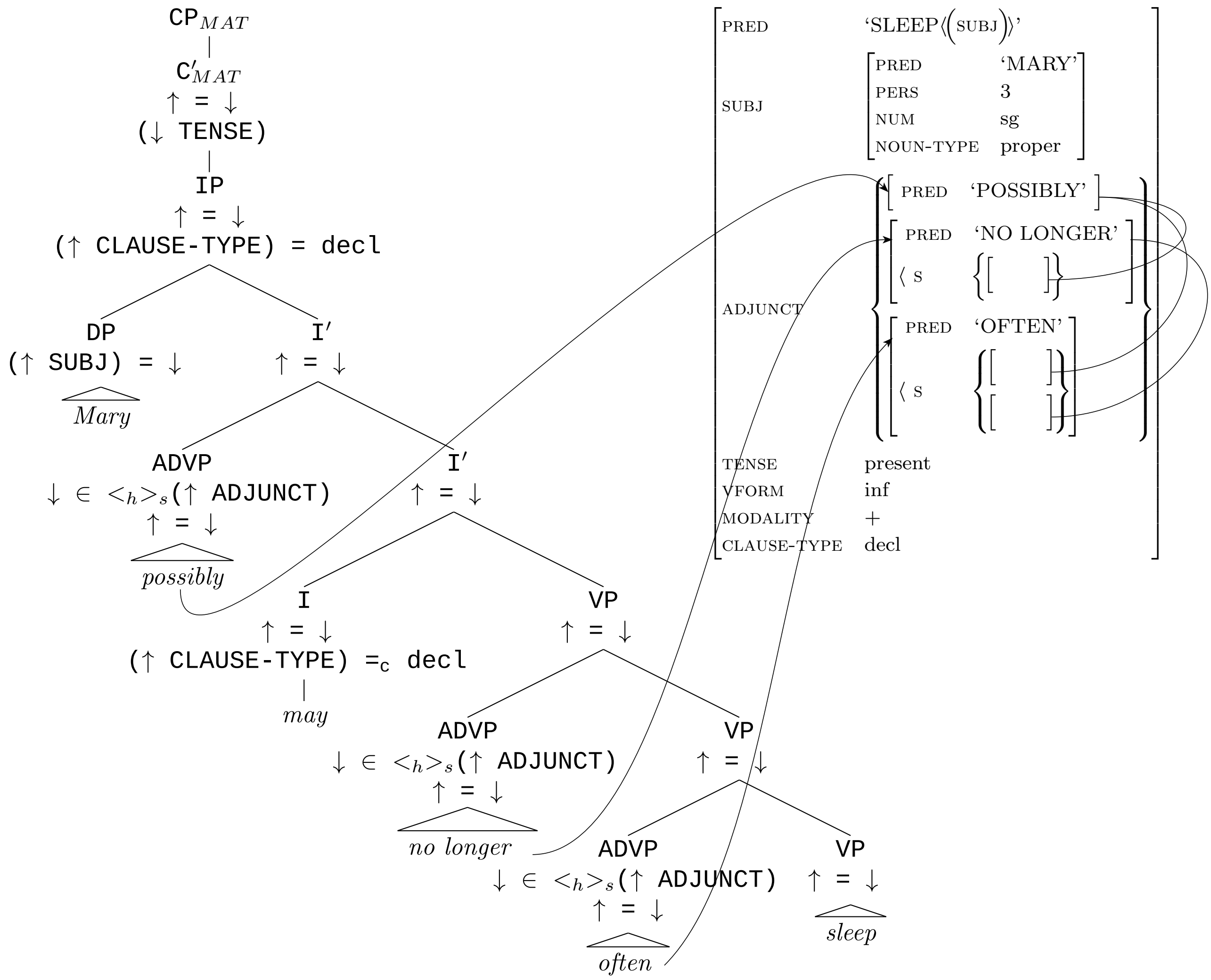
(89) *scope-relational set membership*
$f$ is a scope-relational member of $g$ ($f \in <_h>_s g$) if and only if $f \in g$ and for all other elements $x$ that were also added as scope-relational members to the set $g$, if $f <_h x$, then $f >_s x$.

Scope-relational set membership can then be used to add members to an adjunct set with two rules adjoining an adverb phrase to `I'` and `VP` respectively. These rules are shown in (90).

| | | | |
|---|---|---|---|
| (90) | a. `I'` $\rightarrow$ | `ADVP` <br> $\downarrow \in <_h >_s$ `(`$\uparrow$` ADJUNCT)` | `I'` <br> $\uparrow = \downarrow$ |
| | b. `VP` $\rightarrow$ | `ADVP` <br> $\downarrow \in <_h >_s$ `(`$\uparrow$` ADJUNCT)` | `VP` <br> $\uparrow = \downarrow$ |

The model can now parse sentences with any number of adverbs immediately before and after an auxiliary and represent their relative scope explicitly in the f-structure. This is illustrated for the example *Mary possibly may no longer often sleep* in (91).

(91) Mary possibly may no longer often sleep.




The parse of this sentence proceeds in the same way as for the previous examples. A tense feature is required by the finite matrix clause and is valued as present by the finite verb; a declarative clause type feature is introduced in the IP-layer in the absence of subject-auxiliary inversion; the modal selects an infinitive via a VFORM feature etc. What is new in this example is the presence of one ADVP adjoined to I′ and two ADVPs adjoined to VP. By virtue of a constraint involving the scope-relational set membership operator, they map onto elements of an adjunct set, as shown by the arrows. In addition, every function of that set (e.g., 'possibly') that precedes elements of the same set in the syntactic tree (e.g., 'no longer' and 'often') is unified with a function in these elements' sets of functions that scope over them (labeled $<_s$), as shown by the unification lines in the f-structure. In that way, the functional structure contains the information that *possibly* scopes over *no longer* as well as *often*, and that *no longer* scopes over *often*. Hence, my model can convert relatively informal conditions on scope in syntactic structures (say, "c-command regulates scopal relations") into information contained in actual linguistic representations. As before, the model produces one unique solution for the parse.

Note that my model allows in principle for the generation of an infinite number of adverbs in front of and after auxiliaries. I will not restrict my competence-based model in any way to account for the fact that sentences with a large number of adverbs can be severely degraded. Instead, the degradation of sentences with very many adverbs, such as *Mary thus certainly nevertheless has perhaps usually again seen someone*, is assumed to follow from the same memory and processing limitations as, for instance, the oddity of sentences with multiple center embedding, like *The rat the cat the dog chased killed ate the cheese*, etc. I have nothing to say about how such processing effects arise exactly. Likewise, I will leave for further research the complicated question of how exactly the scope sets should be allowed to manipulate the meaning constructors on adverbs in order to yield one unambiguous logical formula indicating their correct scopal relations during the process of semantic composition (for a starting point, see Andrews 2007).

My grammar fragment is now able to parse the adverbs relevant here in pre- and post-auxiliary position. There is, however, a closely-related phenomenon that has not been implemented in my model yet. This phenomenon concerns the three English quantifiers *all*, *both* and *each* when they occur to the right of a local subject DP they modify and is commonly referred to as 'quantifier float.' There is a vast literature on the topic, which I will not review here because of space constraints (for an overview, see Bobaljik 2003).

What is important for the purpose of this chapter is that floating quantifiers essentially pattern like adverbs, which is why I discuss both topics under the same section heading. Firstly, and most importantly, floating quantifiers have the same distribution as adverbs relative to verbs and can therefore also serve as diagnostics for auxiliary and main verb positions. Just like adverbs, floating quantifiers can be placed before, (92a), but not after finite main verbs, (92b), as well as after auxiliaries, (92c), and marginally even in front of them, (92d) (compare to (62) for the same paradigm with adverbs). As with adverbs, the pre-auxiliary position of floating quantifiers is dispreferred. This becomes clear from judgments provided by native speakers and from the fact that some linguists explicitly downgrade such sentences (for example with one question mark in Osborne 2013). Thus, floating quantifiers may actually reveal the difference between positions of auxiliaries and full lexical verbs somewhat more accurately than adverbs.

(92) a. John and Mary **both** visited Spain.
b. * John and Mary visited **both** Spain.
c. John and Mary <u>have</u> **both** visited Spain.
d. $^{ok}$ John and Mary **both** <u>have</u> visited Spain.

Secondly, ellipsis is not licensed after a sentence final adverb, (93a), and neither can ellipsis occur immediately after a floating quantifier, (93b) (Sag 1978).

(93) a. My brothers have nevertheless married young, and my sisters have (*nevertheless), too.
b. My brothers have all married young, and my sisters have (*all), too.

Finally, both adverbs and floating quantifiers can directly reflect their scoping relations with respect to other adverbs in the surface. Hence, neither induce scope ambiguities in pre- and post-auxiliary position in such cases. The matter is actually substantially complicated by certain interactions between floating quantifiers and negation, genericity and other facts (for details, see e.g., Dowty and Brodi 1984). In the example below, floating *all* and the adverb *always* occur before or after another adverb, *never*, and their positions are directly indicative of their scope relation to that adverb.

(94) a. The arrows **all** will never hit the target.
$\forall >$ *never* (Not a single arrow will ever hit the target)
**never* $> \forall$ (Not all arrows will hit the target, but perhaps some)

b. The arrows will never **all** hit the target.
$*\forall >$ *never* (Not a single arrow will ever hit the target)
*never* $> \forall$ (Not all arrows will hit the target, but perhaps some)

c. The arrows **always** will never hit the target.
*always* > *never* (Not a single arrow will ever hit the target)
**never* > *always* (Arrows will not always hit the target, but perhaps sometimes)

d. The arrows will never **always** hit the target.
**always* > *never* (Not a single arrow will ever hit the target)
*never* > *always* (Arrows will not always hit the target, but perhaps sometimes)

Having shown that adverbs and floating quantifiers distribute alike, I will now turn to their most salient difference, the fact that floating quantifiers modify a subject whereas the adverbs considered here are (at least at first sight) propositional modifiers. The core observation is, of course, that sentences with quantifiers in front of the subject DP over which they quantify are (largely) logically equivalent to sentences where such a quantifier floats to the right of its subject DP.

(95) **All** the knights have slain the dragon.
$\approx$ The knights have **all** slain the dragon.

Linguists have developed two different proposals for the mechanism by which this logical equivalence is achieved. The first analysis is best characterized as a raising approach. In transformational terms, the quantifier and the subject DP it quantifies over are first merged together in an appropriate external argument position and the nominal constituent undergoes leftward A-movement, leaving the quantifier stranded in a lower position (most famously, Sportiche 1988). This approach focuses on the logical similarity between sentences with adnominal and floating quantifiers.

(96) [The knights]$_i$ have [$_{\text{VP}}$ [**all** _$_i$ ] slain the dragon].

The raising analysis can be translated into the LFG framework used here as the annotation of a phrase structure rule that introduces floating quantifiers with a constraint that forces them to unify with a function within a local subject, say the quantification function, QUANT.
The second analysis can be called a binding approach. In derivational terms, it assumes that floating quantifiers are base-generated as anaphors and contain some kind of variable that must be bound by an antecedent according to a binding condition similar to that for reflexives (for one of the most detailed defenses of such an approach, see Bobaljik 1995, chapter 4). This approach focuses on the semantic differences between sentences with adnominal and floating quantifiers.

(97) [The knights]$_i$ have [$_{\text{VP}}$ [**all** PRO$_i$] slain the dragon].

In LFG, such an approach would entail that the floating quantifier heads its own function, presumably a member of an adjunct set just like adverbs, and contains a complement function (say, an object) headed by PRO, which can then be co-indexed with the subject according to some binding constraint (for a formalization of Hebrew floating quantifiers along these lines in LFG, see Spector 2009).

In the following paragraphs, I will summarize the most important pieces of evidence that have been used in the literature to argue in favor of the raising or binding analyses. The overview will be cursory because the issue is quite complex and because I cannot, at any rate, provide an in-depth answer to this long-standing problem within the short space I have at my disposal here.
(i) Some of the evidence presented in the literature seems to me to be irrelevant to decide between one or the other analysis. For example, derivational models of the raising approach lead to rather blatant movement paradoxes, where the postulated underlying constituents are actually ungrammatical. One illustration is given below, but many more could be adduced.

(98) a. [Mary, John and Bill ]$_i$ were [all _$_i$] happy.
b. * [[All [Mary, John and Bill]]]$_i$ were _$_i$ happy.

Furthermore, floating quantifiers can occur in places where derivational models do not normally posit underlying arguments that can be raised to a subject position, most notably in front of auxiliaries, (99a), and at the same time they do not occur in certain positions where such models do assume relevant arguments to be merged, for instance after main verbs in passives, (99b).

(99) a. *floating quantifier but no underlying argument position*
[The knights]$_i$ [$_{?P}$ [**all** _$_i$] have slain the dragon].

b. *no floating quantifier but underlying argument position*
* [The knights]$_i$ have been [$_{VP}$ slain [**all** _$_i$]].

Within a derivational syntactic framework, such arguments may perhaps favor an analysis that treats floating quantifiers as anaphors. However, in a representational framework like the one used here, they carry relatively little weight. What is at issue is not the distribution of floating quantifiers, but merely the mechanism that establishes a dependency between them and their local subject.

(ii) Various data points that are sometimes mentioned in favor of one or the other analysis seem to me to be very arcane and ultimately inconsequential. I will give just two examples, but many more interesting observations could be added. Firstly, floating quantifiers may have an overt pronominal morpheme in some languages or earlier stages of English, (100), which may support their status as anaphors. It is, however, not clear if such data can readily be transferred to Modern English.

(100) And the four beasts had <u>each of them</u> six wings about him.
*King James Bible*, Revelation 4:8 (c. 1611 A.D.)

Secondly, object-oriented floating quantifiers modifying non-pronominal objects are not generally acceptable, (101a), but may be licensed if certain phrases follow, (101b) (for some discussion, see Maling 1976). The reason for such contrasts may not lie in one unified factor, but could result from an interplay of structural constraints, stress requirements as well as the difference between propositional and non-propositional phrases. Again, it is not clear how exactly such data informs the relevant choice between the raising and binding analyses.

(101) a. * I saw the kids **all**.

b. I gave the kids **all** some candy. (adapted from: Maling 1976: 715, ex. (16a))

(iii) It is sometimes argued that the fact that a binding condition for floating quantifiers can be modeled on one for reflexive anaphors supports the binding analysis (see the discussion in Bobaljik 2003 and references therein). I believe such a line of reasoning to be fallacious, as I will argue now. It is true that the basic requirements for licensing reflexives are the same as those for floating quantifiers. Firstly, Binding Principle A (regulating reflexives) requires that reflexives be bound within a certain binding domain, roughly a phrase that contains both the reflexive and a subject. In the same way, a floating quantifier must occur within a domain that contains both the floating quantifier and a subject, which it must then refer to. In the examples below, the domain wherein the reflexive respectively the floating quantifier must be licensed are indicated with square brackets. Here, this domain contains neither a suitable antecedent for the reflexive nor a suitable subject for the floating quantifier. Both sentences are therefore ungrammatical.

(102) a. * [My friends]$_i$ think that [I have admired **themselves**$_i$.]

b. * [My friends]$_i$ think that [I have **all**$_i$ left.]

Secondly, in a traditional formulation of Binding Principle A, a reflexive must be c-commanded by its antecedent. Similarly, a subject must c-command its floating quantifier. The examples below do not feature adequate c-command relations and are therefore ungrammatical. (Such a constraint would be formulated somewhat differently in most LFG treatments of the topic.)

(103) a. * [The mother of [my friends]$_i$] has admired **themselves**$_i$.

b. * [The mother of [my friends]$_i$] has **all**$_i$ left.

However, Binding Principle A is also designed to deal with much more complex cases of binding, which are not applicable to floating quantifiers. For instance, reflexives can be bound when they are complements of prepositions for example in picture DPs, (104a), or in prepositional arguments, (104b). Reflexives can be bound even if they have topicalized in front of their antecedent, (104c), or when they are subjects of non-finite or small clauses, (104d).

(104) a. [My friends]$_i$ saw pictures of **themselves**$_i$.
b. [My friends]$_i$ listened to **themselves**$_i$.
c. **Themselves**$_i$, (I think) [my friends]$_i$ admired.
d. [My friends]$_i$ expected [ **themselves**$_i$ to win ] / considered [ **themselves**$_i$ lucky ].

None of these positions are available for floating quantifiers. This may follow straightforwardly from phrase structure rules determining where floating quantifiers can be inserted into the syntactic structure.

(105) a. * [My friends]$_i$ saw pictures of **all**$_i$.
b. * [My friends]$_i$ listened to **all**$_i$.
c. * **Both**$_i$ , (I think) [my friends]$_i$ left.
d. * [My friends]$_i$ expected [ **each**$_i$ to win ] / considered [ **each**$_i$ lucky ].

Hence, Binding Principle A is needlessly powerful to account for the distribution of floating quantifiers. It vacuously specifies a whole range of cases of binding for floating quantifiers in places where they cannot occur to begin with. Consequently, the parallelism between Binding Principle A for reflexives and the licensing conditions on floating quantifiers cannot be used to argue that a binding principle is likely required to capture the distribution of both. The dependency between floating quantifiers and their subjects can be captured with much simpler constraints.
(iv) The most important predictive difference between the two analyses concerns the (im)possibility of multiple occurrences of subject-oriented quantifiers. The raising analysis predicts that there can be only one quantifier modifying its local subject. The reason is, trivially, that a function must have exactly one unique head (irrespective of where the head is introduced in the syntactic tree and irrespective of how many places the function is interpreted in). A simple example is shown in (106). The associate of an existential *there* expletive can be placed inside the verb phrase or it can topicalize to the beginning of the sentence, but it cannot occur in both positions at the same time because the associate function would be doubly-headed.

(106) *Raising: Multiple heads of the same function are impossible*
a. There was [a child] in the garden.
b. [A child], there was in the garden.
c. * [A child], there was [a child] in the garden.

In contrast, the binding analysis predicts that there can be, in principle, an infinite number of quantifiers that depend on the same subject. That is because one antecedent can bind any number of anaphors as long as they meet the right binding conditions. This is illustrated in (107), where one antecedent binds a large number of possessives and reflexives.

(107) *Binding: Multiple anaphors bound by the same antecedent are possible*
[My friends]$_i$ presented their$_i$ pictures of themselves$_i$ to themselves$_i$ in their$_i$ mirrors.

The question is thus whether or not multiple quantifiers can establish a dependency with the same local subject. In general, the answer to this question appears to be negative. However, the matter is far from being clear-cut. Firstly, my native speaker informants reject many sentences with two floating quantifiers, (108a). At the same time, there are some cases of quantifier stacking, for example with *each*, (108b). I assume that such counterexamples are somehow marginal - perhaps they involve a different form of *each* indicating some kind of distributivity - and should not be generated by a rule for floating quantifiers.

(108) a. * My sisters have **both all** collected a flower.
b. Ross McCormack and Rudy Gestede have **both each** scored 7 goals in 7 games
*Sportdec* on Twitter, 11 September 2016

Secondly, adnominal quantifiers in subject position are normally incompatible with co-occurring floating quantifiers, (109a), but, again, there appear to be counterexamples, (109b). These counterexamples, too, may be structurally different from those that I want to generate with a designated rule for floating quantifiers. Specifically, example (109b) seems to involve a form of *all* that expresses a sense of togetherness rather than quantification in conjunction with the predicate *have left in one car*. If the predicate is changed to, say, *have won*, the sentence becomes ungrammatical.

(109) a. * {**All** / **Most** / **Some**} mathematicians have **both** solved seven problems.

b. **Some** of the students might **all** have left in one car. (from: Bobaljik 1995: 225, ex. (48a))

Thirdly, most of my native speaker informants dislike identical floating quantifiers in different positions relative to multiple auxiliaries in one clause (110a), and, less so, in cases where they refer to different instances of a chain of subjects in subject-to-subject raising contexts, (110b). However, the generalization is once again not unproblematic since some native speakers actually appear to find those sentences perfectly acceptable, perhaps to express emphasis. I therefore mark these examples with two question marks rather than as downright ungrammatical.

(110) a. ?? My friends may **both** have **both** been singing.

b. ?? My friends **all** seemed to **all** win.

As a rule of thumb, then, only one quantifier can quantify over the local subject. However, it must be stressed that the evidence is not at all unequivocal. The fact that quantifier stacking is usually disallowed merely reflects the most generic generalization that can be derived from the data. An adequate model of floating quantifiers would have to deal with the numerous exceptions in a much more nuanced manner.

Having reviewed the most important evidence for and against the raising and binding accounts, I will now develop the rules necessary to parse simple sentences involving floating quantifiers. I decided to implement a version of the raising account in my model. This is not to say that I am wholeheartedly convinced that this is the correct approach since the evidence is meager and ambiguous, as I have just shown. Rather I opted for this approach because I think it is both somewhat simpler to implement and more widely assumed in the field and because the general (though not exceptionless) ban on multiple quantifiers accords better with the raising than the binding analysis. Specifically, I propose the phrase structure rules in (111) for parsing sentences with floating quantifiers. They allow a floating quantifier phrase, abbreviated `FQP`, to adjoin to `I'`, (111a) or `VP`, (111b). This captures in a straightforward way the initial observation that floating quantifiers distribute like adverbs. Floating quantifier phrases can only be re-written to elements of category floating quantifier, `FQ`, (111c). The rules are annotated with a functional constraint that forces the floating quantifier to unify with the local subject's quantification function, `(↑ SUBJ QUANT) = ↓`. This constraint is in essence a version of the raising analysis. It correctly captures the locality constraints described above - a floating quantifier can only modify a local subject and not one of a higher clause (see (102b)), and it can only refer to quantification of the subject function directly and not to quantification of a function embedded within it, like a possessor (see (103b)). The constraint also entails the presence of only one quantifier since the quantification function must have one and only one head (see (108) - (110)).

| | | | |
|---|---|---|---|
| (111) a. | `I'` → | `FQP`<br>`(↑ SUBJ QUANT) = ↓` | `I'`<br>`↑ = ↓` |
| b. | `VP` → | `FQP`<br>`(↑ SUBJ QUANT) = ↓` | `VP`<br>`↑ = ↓` |
| c. | `FQP` → | `FQ`<br>`↑ = ↓` | |

The lexical entries for floating quantifiers can be specified with relatively sophisticated constraints to assure the presence of an adequate subject. I will illustrate such lexical constraints for the quantifier *all*. It can occur with a large number of different subjects: plural, but not singular, personal pronouns, (112a), plural common count nouns when they are definite, (112b), but not when they are bare, (112c), definite mass nouns, (112d) (but *both* and *each* cannot occur with definite mass nouns, (112e)), and conjunct sets with at least 3 members, (112f), but not fewer than 3 members, (112g), and not with disjuncts, (112h), and so on and so forth.

(112) a. {We / * I } could **all** have slept on the floor.

b. The apples have **all** been eaten.

c. * Apples have **all** been eaten.

d. The data has **all** been lost.

e. * The data has **each** been lost.

f. John, Bill and Mary will **all** come to the party.

g. * John and Bill will **all** come to the party.

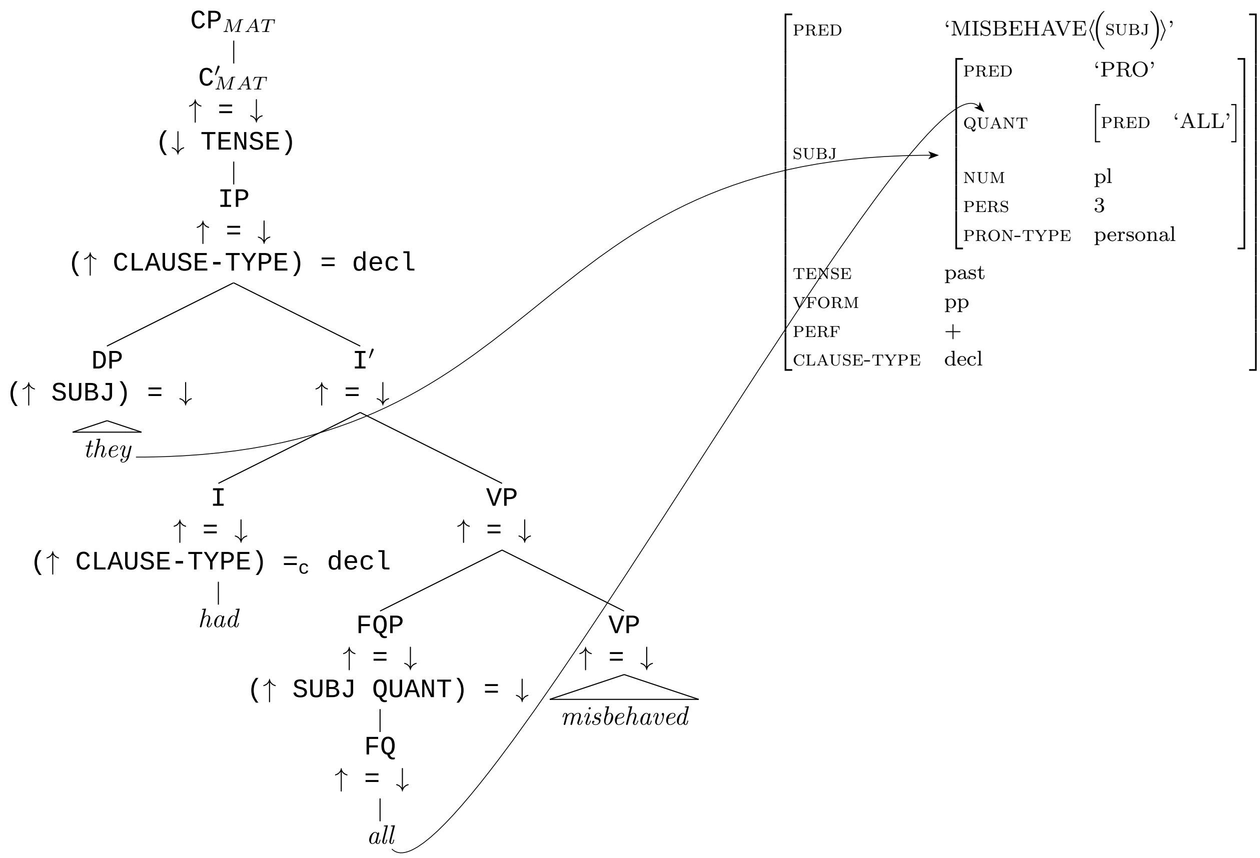
h. * John, Bill or Mary will **all** come to the party.

A lexicon entry for *all* is presented below in (113). It lists a set of lexical constraints permitting only certain kinds of subjects, namely plural personal pronouns, plural definite common count nouns, definite mass nouns, and conjunct set with more than 2 members, in that order. Several more would have to be added to reach a more comprehensive coverage. Such a long list of subject-regulating constraints not only seems inelegant but may also miss important semantic generalizations. Further, the constraint that restricts the cardinality of the function containing quantification headed by *all* to more than 2 is not standardly assumed in LFG. It would therefore probably be better if the co-occurrence restrictions between a subject and a floating quantifier could be captured, at least in part, by the semantic denotations of the interacting semantic resources. Here, I present a purely feature-based implementation of the restrictions imposed on a floating quantifier's subject for illustrative purposes and in order to develop an actual workable toy grammar.

(113) *all* FQ (↑ PRED)= 'ALL'
{ ((QUANT ↑) NUM)=$_c$ pl
((QUANT ↑) PRON-TYPE)=$_c$ personal
| ((QUANT ↑) NUM)=$_c$ pl
((QUANT ↑) DEF)=$_c$ +
((QUANT ↑) NOUN-TYPE)=$_c$ common-count
| ((QUANT ↑) DEF)=$_c$ +
((QUANT ↑) NOUN-TYPE)=$_c$ mass
| |(QUANT ↑)| >2
((QUANT ↑) COORD-FORM)=$_c$ and }

My model can now parse a number of sentences with floating quantifiers, such as *They had all misbehaved.*

(114) They had all misbehaved.




The subject DP of this sentence is the pronoun *they*, which introduces the features 3$^{rd}$ person plural and personal pronoun type into the subject function, as emphasized by an arrow ranging from *they* to its f-structure. The set of features in the subject function are compatible with the constraints imposed on it by the floating quantifier *all*, specifically the first option in (113), which requires the presence of a plural personal pronoun subject. The floating quantifier *all* can therefore be introduced into the structure by the rules in (111), either before or after the auxiliary position. Here, it occurs after the auxiliary *had*, adjoined to VP. The rule introducing a floating quantifier forces it to head a quantification function associated with the subject via the constraint ($\uparrow$ SUBJ QUANT) = $\downarrow$. As a consequence, the word *all* quantifies over the subject, as desired. The discontinuous unification of the two parts of the subject function is indicated by an arrow leading from the word *all* to the quantification function QUANT contained in the subject f-structure. The rest of the analysis proceeds as before, yielding one unique solution for the parse.

Just as for my analysis of adverbs, a great many open questions remain for the proper treatment of floating quantifiers. Firstly, my model can generate sentences with floating quantifiers along with several other elements before an auxiliary, which may sometimes be considered unnatural or even ungrammatical, for instance *My friends all probably will leave*, and again, I attribute this to processing effects and will not restrict my grammar model any further to exclude such structures. Moreover, a satisfying analysis of floating quantifiers must ultimately provide formal semantic denotations for them. Only then can one discern different uses of floating quantifiers, such as the use of *all* to express quantification or togetherness, explore meaning differences of floating quantifiers in pre- and post-auxiliary position, explain exceptions to the ban on multiple quantifier occurrences, and understand what it might actually mean to quantify over proper names or pronouns etc. Finally, I did not extend the scope mechanism developed above to floating quantifiers and it remains difficult to make meaning constructors sensitive to surface scope for the purpose of semantic composition.

Despite these shortcomings, I hope to have built a toy grammar that encapsulates at least the essence of customary views regarding the syntax of sentential adverbs and floating quantifiers. In the next section, I will discuss one last context that informs the syntactic difference between auxiliaries and main verbs, namely verb phrase ellipsis structures.

#### 2.2.1.6 Verb Phrase Ellipsis

Verb phrase ellipsis (VPE) structures are constructions involving an antecedent clause with a main verb and all of its arguments, called the source clause, and a sentence with a subject and an auxiliary, but no other verb phrase material like a main verb and its dependents, called the target clause. The meaning of the source clause determines the meaning of the target clause. VPE comes in a number of different subclasses, such as deletion under positive polarity with *too*, partial questions and tag questions, or cases of inverted polarity after negation etc. If a source clause includes an auxiliary, the target clause can occur without a verb phrase after the same auxiliary, (115a). If a source clause does not include an auxiliary, however, verb phrase ellipsis after the full lexical verb is ungrammatical, (115b), and requires the dummy auxiliary *do* instead, (115c). Finally, if a source clause involves an auxiliary, the target clause is degraded, if not ungrammatical, if it uses *do*-support instead of the contextually salient source auxiliary, (115d). Hence, VPE is another context that is sensitive to the distinction between auxiliaries and main verbs. In the examples below, auxiliaries are underlined and a delta represents the verb phrase material understood from the source clause.

(115) a. You <u>were</u> spilling the beans too early.
And they <u>were</u> Δ, too. / Or <u>were</u> you Δ? / But I <u>was</u> not Δ.

b. You spilled the beans too early.
* And they spilled Δ, too. / * Or spilled you Δ? / * But I spilled not Δ.

c. You spilled the beans too early.
And they <u>did</u> Δ, too. / Or <u>did</u> you Δ? / But I <u>did</u> not Δ.

d. You <u>were</u> spilling the beans too early.
?? And they <u>did</u> Δ, too. / ?? Or <u>did</u> you Δ? / ?? But I <u>did</u> not Δ.

Where the last section was struggling to present a model of adverbs and floating quantifiers reflecting the current consensus in the field, the present discussion of VPE cannot possibly hope to result in anything near an uncontentious understanding of the topic, not even for at least a sizable number of linguists. There are simply too many different approaches to VPE (for an overview over different analyses of the phenomenon, see Merchant 2016), and too many related phenomena, such as VPE after infinitival *to* or multiple auxiliaries, N-bar deletion, stripping, gapping, fragment answers, antecedent contained deletions, etc. (for an overview, see Johnson 2001). My analysis will therefore be very cursory and focus only on the most essential aspects needed to parse basic cases of VPE within the model assumptions developed here.

There are two radically different approaches to the syntax of ellipsis. Structural approaches assume that there is unpronounced syntactic structure in the ellipsis site. I will restrict myself to structural accounts that assume that VPE involves complete phrase markers from which the meaning of the structure can be inferred and that the syntactic representation subsequently gets deleted yielding an ellipsis site. The empty syntactic structure may thus contain deleted lexical and functional material as well as traces or copies of moved elements. It seems fair to say that the majority of theories of verb phrase ellipsis have been formulated within this structural paradigm.

Non-structural approaches, in contrast, assume that VPE does not involve deleted syntactic material, but instead develop semantic devices that can generate meanings in the absence of syntactic structure. At the moment, only a relatively small number of linguists subscribe to the non-structural paradigm. However, LFG does not incorporate transformations, which operate on trees to create trees, like deletion transformations (whether at PF or elsewhere is immaterial). Theories of VPE formulated within that framework will thus quite naturally lean towards a non-structural approach. I will sketch the bare essentials of a non-structural approach to verb phrase ellipsis in LFG in the subsequent sections.

The starting point of my discussion concerns the idea that the syntactic model should be able to blindly generate verb phrase ellipsis structures. In my view, a string with elided material, say, *Bill is* $\Delta$, should always be regarded as syntactically well-formed irrespective of the context it is used in. The syntactic component of the grammar does not know if an appropriate source clause is available to tolerate VPE in discourse. Syntax must simply be able to provide verb phrase ellipsis constructions independent of the pragmatic context. In other words, I propose a radical separation of the problem of syntactic licensing of ellipsis structures, which is part of grammatical knowledge, from the problem of identifying suitable antecedent clauses, which is an inference process or pragmatic act (see also the logic of positive *do*-support structures in section 2.2.1.2).

Focusing on the syntactic structure of VPE, then, I suggest that for every phrase structure rule involving a head on the left hand side and a `VP` node on the right hand side, there is a second phrase structure rule that truncates the `VP` category to the empty string, $\epsilon$. This is a direct implementation of the classical observation that VPE targets `VP`s, or, alternatively, that VPE is licensed by an auxiliary of category `I`. My model includes three such phrase structure rules: (i) There is a rule that re-writes `I'` to `I` and `VP` and constrains the clause type to be declarative, (10a). (ii) The second rule re-writes `I'` to a negatively contracted auxiliary of type `I+NEG` and `VP` in declarative contexts , (25). (iii) The final rule re-writes `NEGP` to `NEG` and `VP` without further restrictions, (35b). Hence, there will be three new rules, in which the `VP` node is replaced by the empty string $\epsilon$ while all the other constraints are left intact. Additional rules would be required to model the interaction of ellipsis with subject-auxiliary inversion. The empty string effectively represents nothing in the syntax - it counts neither as a node in the syntactic tree, nor as a pro-form, nor as deleted material. It is included mainly to illustrate the parallelism with the non-elided phrase structure rules.

Because of the technical way my grammar fragment is set up, the verb phrase ellipsis rules must be associated with two kinds of constraints. Both of them could likely be avoided completely in a more particularized model. The first constraint determines the subcategorization properties of the missing main verb. My model takes a naive view on valency, in which a predicate directly lists the arguments it requires. More recent deliberations on argument structure in LFG have shown that it is possible to endow a predicate with a stripped-down `PRED` feature that merely ensures that its f-structure is uniquely headed and does not contain information on subcategorization (see among several others, Asudeh et al. 2014, Asudeh 2012). The phrase structure rules in my model, however, require the introduction of a traditional `PRED` feature in order to permit a solution for the parse. The missing verb is represented as an underspecified predicate, $P$. Its subcategorization frame is selected from a set of all relevant subcategorization frames for main verbs. In the phrase structure rules below, this is illustrated by a small set of possible subcategorization frames - an intransitive verb (selecting a subject, `SUBJ`), and a transitive verb (selecting a subject and an object, `SUBJ`, `OBJ`). The different valencies will be paired with corresponding denotations and meaning constructors. The second constraint ensures that the selectional requirements imposed by the auxiliary are satisfied. As explained earlier, my model includes a somewhat inelegant `VFORM` feature with values such as present participle, infinitive etc., which is introduced by the main verb, and simultaneously checked by an auxiliary, such as progressive *be*, a modal etc. In a more comprehensive grammar the `VFORM` feature could be dispensed with, for example by means of typed `VP` nodes. Since all kinds of auxiliaries can engage in verb phrase ellipsis, the elided verb phrase must permit the unrestricted introduction of whatever verb form is required from the set of all possible values of the `VFORM` feature. This is formalized in the phrase structure rules below as a disjunctive set of all possible pairs of the `VFORM` attribute and its values. The resulting constraints will be added to the three `VP`-sensitive rules mentioned above. However, for reasons of perspicuity, I will display the constraints on the empty string on the last rule only.

(116) a. `I′` → `I` ($\uparrow = \downarrow$, `(↑ CLAUSE-TYPE)` $=_c$ `decl`) $\epsilon$ ($\uparrow = \downarrow$)

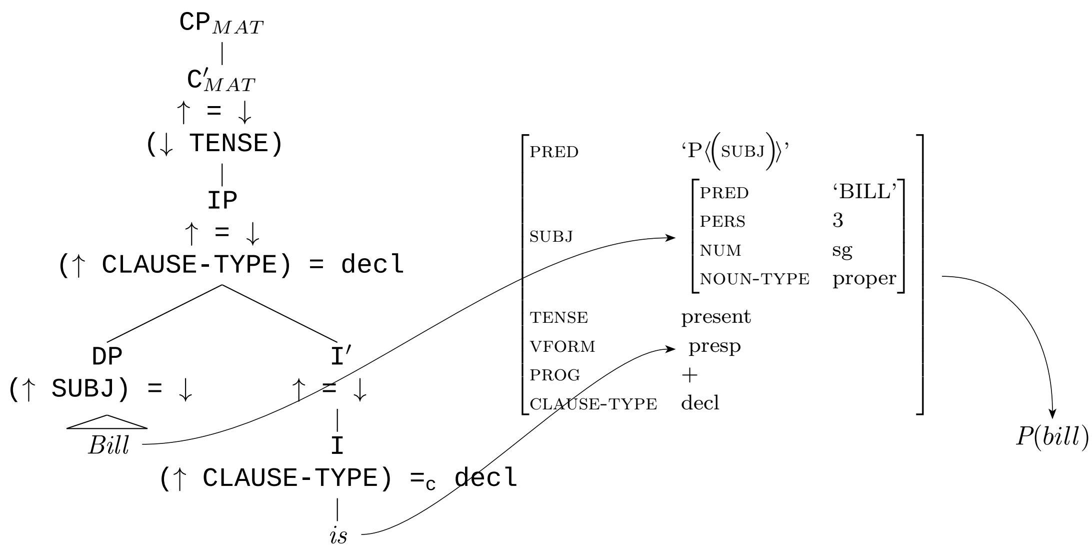
b. `I′` → `I+NEG` ($\uparrow = \downarrow$, `(↑ CLAUSE-TYPE)` $=_c$ `decl`) $\epsilon$ ($\uparrow = \downarrow$)

c. `NEGP` → `NEG` ($\uparrow = \downarrow$) $\epsilon$

```
↑ = ↓
{ (↑ PRED) = 'P <(↑ SUBJ)>'
  λx.P(x): (↑ SUBJ)σ ⊸ ↑σ
| (↑ PRED) = 'P <(↑ SUBJ) (↑ OBJ)>'
  λx.λy.P(x)(y): (↑ SUBJ)σ ⊸ [ (↑ OBJ)σ ⊸ ↑σ ]
| ...   }
{ (↑ VFORM) = inf
| (↑ VFORM) = presp
| (↑ VFORM) = pp
| (↑ VFORM) = fin }
```

My model can now parse ellipsis structures out of context and provide a semantic representation for them. An example is shown below for the VPE structure *Bill is* Δ.

(117) Bill is Δ.




The model parses *Bill* as the subject of this sentence. There is a progressive auxiliary, *is*, which is parsed as the head `I` and requires the presence of a present participle feature, `VFORM` $=_c$ `presp`, in its f-structure. These two parsing aspects are illustrated above with the arrows leading from the lexical items in the syntactic tree to their f-structural values. Since there is no verb phrase material in this structure, the sentence must be parsed with a VPE rule, specifically, the rule in (116a), which truncates the syntactic structure at the `I`-position. An option from the disjunctive set of subcategorization frames and verb form features is picked in a way that enables the parser to arrive at a solution for the parse. To be precise, the missing predicate has the valency of an intransitive verb to be able to apply to the only grammatical function, subject, in this structure, `'P <(↑ SUBJ)>'`, and it takes the form of a present participle to satisfy the form constraint imposed by the auxiliary, `VFORM = presp`. Thus, the parser produces one unique solution for this sentence. Further, the meaning constructors on the lexical items and phrase structure rules allow for the composition of exactly one logical formula representing the meaning of the sentence. The semantic composition, which is trivial and therefore not shown above, yields a formula that predicates some underspecified property, $P$, of the subject, *Bill*, i.e., $P(bill)$.

The model also correctly rejects VPE structures that terminate in phrasal elements such as adverbs or floating quantifiers (e.g., Engels 2010). This follows simply from the non-availability of relevant phrase structure rules that could truncate the syntactic tree after such categories.

(118) * Bill [$_{\mathrm{I'}}$ is [$_{???}$ often Δ ]].

How is the meaning of the underspecified target predicate computed on the basis of an antecedent? The most widely used algorithm to solve this problem within LFG is a version of the equational analysis proposed by Dalrymple et al. (1991). In what follows, I will paraphrase their account in my own words as a four-step process.

First, a source clause must be identified in the form of a logical formula representing its meaning. As stated above, this identification is a pragmatic act over discourses and cannot be achieved on purely syntactic grounds alone. Correspondingly, relatively few conditions on antecedent identification have been discussed in the literature (for a summary, see Johnson 2001: 447-455). In general, the task does not seem to be extremely difficult. For example, source sentences can be automatically identified with good precision in parsed corpora (Hardt 1997, Nielsen 2004). I will not attempt to formalize restrictions on possible and impossible positions of source clauses any further but merely presuppose that they can be identified accurately. An example is shown in (119), where the clause on the first line and its semantic representation has been identified as the source for the clause on the second line with its respective semantic representation as the target.

(119) *(1) Identification of source and target*
Mary isn't reading *De Amicitia.* $\rightsquigarrow \neg(read(mary, de\text{-}amicitia))$
Bill is Δ. $\rightsquigarrow P(bill)$

Secondly, the ellipsis resolution algorithm assumes that there is a certain degree of parallel structuring between the target and source sentences such that speakers are able to identify certain elements in the target as parallel to certain elements in the source. The specific process by which this identification occurs is a matter of some debate (for cases of establishing parallel elements on non-syntactic grounds, see Dalrymple et al. 1991, section 5; for formal considerations, see Hobbs and Kehler 1997). In the simplest terms, one can define parallel elements on purely syntactic grounds - as parallel grammatical functions or features that are present in identical contexts in both the target and source sentences. In the present example, there would thus only be one relevant parallel element, namely the subject function - the subject *Bill* in the target is parallel to the subject *Mary* in the source. Indeed, VPE structures typically, by definition maybe even necessarily, involve a parallelism between the subject functions of source and target clauses.
The resolution algorithm further assumes that it is the property that applies to the parallel element(s) in the source that is in some sense close to the underspecified property, $P$, of the target. Specifically, one can identify some predicate in the source sentence as the basis for the underspecified property, $P$, if its argument (or arguments) correspond to an element (or elements) identified as parallel between the source and target clauses. In the present example, the function that applies to the sole parallel element in the source, i.e., the subject *Mary*, is the predicate *read*.
Next, it is possible to recover via some semantic computation the exact meaning of the underspecified target property $P$. This semantic computation requires the establishment of a certain equation, which I will call the 'parallelism equation.' This parallelism equation essentially holds between (i) the underspecified property $P$ applied to the parallel element(s) of the source on the one hand and (ii) the interpretation of the corresponding predicate and everything in its scope in the original source clause on the other hand. In the current example, the parallelism equation must relate the underspecified predicate $P$, predicated of *Mary*, to the source predicate *read* and its arguments, *Mary* and *De Amicitia*. Furthermore, every semantic element that scopes over the source predicate and its arguments must remain intact and also scope over the corresponding underspecified property, $P$, applied to the parallel element(s). In the present case, negation scopes over the source predicate *read* and so must also scope over the underspecified predicate $P$. One thus arrives at the following parallelism equation.

(120) *(2) Determination of the parallelism equation*
$\neg P(mary) = \neg(read(mary, de\text{-}amicitia))$

In a third step, the parallelism equation is solved to arrive at the meaning of the underspecified predicate, $P$. The specific mechanism to solve the parallelism equation can be quite complex and involves Huet's higher-order unification as a means to completely enumerate all possible solutions as well as ignoring certain potential solutions for which replacement leaves so-called primary occurrences in the result (for details, see Dalrymple et al. 1991, Crouch 1999). Here the solution for $P$ denotes the set of individuals that read *De Amicitia*, (121).

(121) *(3) Solution of the parallelism equation for P*
$\neg P(mary) = \neg(read(mary, de\text{-}amicitia))$
$P(mary) = read(mary, de\text{-}amicitia)$
$P = \lambda x.read(x, de\text{-}amicitia)$ (by higher-order unification / second-order matching)

Finally, the meaning of the VPE structure can be computed in a straightforward way by substituting the resultant property for the underspecified VPE predicate in the target. Here, the solution for the VPE structure yields the proposition that Bill is reading *De Amicitia*, which correctly corresponds to intuitions about the meaning of the target clause under investigation. Thus, the meaning of ellipsis structures becomes computable on the basis of semantic representations alone, (122).

(122) *(4) Substitution of the solution of P for the underspecified predicate in the target*
$P = \lambda x.read(x, de\text{-}amicitia)$
$P(bill) = \lambda x.read(x, de\text{-}amicitia)(bill)$
$= read(bill, de\text{-}amicitia)$

Appendix B includes further information on my approach to verb phrase ellipsis in LFG. It defends my non-structural model against some common attacks on such a conception of ellipsis phenomena. I refer the interested reader to the discussion at this place. Here, I will not deal with verb phrase ellipsis in any more detail. Instead, I will merely express my hope that, while the specifics of my analysis of VPE are controversial and cannot be translated straightforwardly into all alternative frameworks, at least the most basic aspects of my model could be accepted by almost all linguists: VPE targets VP-nodes in declaratives and the meaning of the ellipsis is recovered based on a relatively parallel antecedent clause. I will conclude this part by summarizing the grammar model thus developed.

#### 2.2.1.7 Summary

The preceding sections laid out a formal grammar model of Modern English clause structure.[4] In constructing this model, I pursued one main goal: I aimed for as high as possible a degree of theoretical concurrence. The general tenets of my model were extrapolated from several widely used analyses. It is also possible to translate the general aspects of my model into most other alternative grammatical frameworks (with some exceptions, such as adverb placement, which was not implemented as a hierarchy of projections, or ellipsis, which was not modeled with empty syntactic structure). As far as formal linguistic models go, then, I believe the broad outline of my characterization of Modern English can be regarded as relatively uncontroversial. As a consequence, it should also be quite uncontentious that the four core patterns discussed - negation, inversion, adverb placement and verb phrase ellipsis - do not just form independent constructions, but are related in a criss-crossing network of syntactic primitives. These phenomena can therefore all diagnose the fundamental distinction between auxiliaries and main verbs in an isomorphic manner.

[4]The grammar model has been implemented as an LFG toy grammar that can be run on the Xerox Linguistic Environment (XLE) grammar engineering platform. This makes it easy to test the model predictions regarding the acceptance and rejection of certain strings and underlines its mathematical explicitness. The grammar can also be viewed with any word editor. It is located on the Dissertation DVD under `Chapter 2 - Possessive Have/02 Toy Grammar/Ch2PossHave.lfg`.

### 2.2.2 The Change

I will now formally pinpoint the way in which late Modern American English possessive *have* changes. Two empirical predictions will be derived from this formalization: First, the realization of possessive *have* should change in a specific direction in the syntactic contexts outlined in the previous sections. Secondly, the change should proceed at the same rate in all of these contexts, i.e., the change in possessive *have* should exhibit a constant rate effect. Finally, I shall show that the change in possessive *have* features several advantageous characteristics that make it an almost ideal test case for the investigation of the Constant Rate Hypothesis.

#### 2.2.2.1 How Possessive *have* Changes

In order to explain how possessive *have* changes during the course of the late Modern period, it is necessary to add a lexicon entry for this item to the grammar model. The featural make-up of the word can be formalized as follows.

(i) Possessive *have* selects for two arguments. Within the assumptions of the LFG framework used here, one must explicitly identify the kinds of grammatical functions selected for. The higher argument indisputably gets mapped onto a subject function. The lower argument, however, has a more obscure status. There are two pieces of evidence suggesting that it is implausible to regard possessive *have* as an ordinary transitive verb that selects for an object function. First, objects can passivize, but the lower argument of possessive *have* cannot normally become the subject of a passive clause. This is true in cases with stative, inalienable, (123a), as well as alienable possession, (123b). Passives involving more dynamic senses of possessive *have* are somewhat more acceptable, (123c). Solely lexicalized passive idioms projected from possessive *have* sound perfectly natural, the most common one relating to jocular expressions of 'having a good time,' (123d).

(123) a. * Blue eyes were had by Mary.
b. * A new car was had by Mary.
c. ? the context in which the discussion was had
d. Fun was had by all.

Secondly, possessive *have* can express an exceptionally wide range of fuzzy possession relations. This is not the case for ordinary transitive verbs, most importantly possessive *have*'s close synonyms *to possess* and *to own*, whose lexical semantics are not so indeterminate. The relations of possessive *have* can range from clearly stative, concrete possession, (124a), over kinship possession, (124b), less stative, abstract possession, (124c), and light verb uses, where the meaning of *have* and its indefinite mono-morphemic complement resemble the meaning expressed by the complement alone, (124d), to idiomatized expressions with bare noun complements (124e). For the first example, the verbs *have*, *possess* and *own* are all possible. For the second, *possess* and *own* convey an unusual sense of servitude. For the subsequent examples, *possess* and *own* become increasingly incompatible with the more dynamic sense of the verb phrases.

(124) a. Mary { has | possesses | owns } a videotape.
b. Mary { had | ? possessed | ? owned } a brother.
c. Mary { has | ? possesses | * owns } a nice personality.
d. Mary { had | * possessed | * owned } a laugh.
e. Mary is { having | * possessing | * owning } lunch.

Work on grammatical relations has frequently carved out a distinction between primary, semantically unrestricted and secondary, semantically restricted functions (for early antecedent ideas, see Fillmore 1968; for Relational Grammar, see Perlmutter and Postal 1983, Dryer 1986; for work in LFG, see Bresnan 1982, more recently Dalrymple and Nikolaeva 2011: 22-28). I will use this distinction to differentiate the complement of possessive *have* from an ordinary object and label it instead as a secondary, semantically restricted object, abbreviated $\mathtt{OBJ}_\theta$. This label is supposed to summarize (though of course not explain) the peculiarities of the internal argument of possessive *have*. First, secondary objects do not show all the core properties associated with objects in a given language. Since passivization is the most important object characteristic in English, and objects of possessive *have* do not readily undergo passivization, it seems warranted to label this grammatical function accordingly. Moreover, secondary objects are semantically restricted. Here, the theta role of possessive *have*'s object function must take on a very distinctive value, albeit wide-ranging, presumably a special kind of theme. I will refer to this theta role simply as the 'possessed element.' Thus, the thematic distinctness of the internal argument of possessive *have* is compatible with its designation as a secondary object as well.

(ii) Next, possessive *have* shows typical verbal characteristics. Its finite forms inflect for tense. The item also agrees with the subject in person and number. These observation can easily be modeled by endowing the lexicon entry of possessive *have* with the corresponding morpho-syntactic features.

(iii) This leaves solely the syntactic category of possessive *have* undiscussed. The word's category status is also the crucial aspect that will be used to model its change. One can simply assume that possessive *have* has undergone a category change from I, the class of Present-Day English auxiliaries, to V, the category of Present-Day English full lexical verbs. Speakers before the 19th century would represent the word only as I, speakers after the 21st century only as V, and speakers during the intermediate period could generate structures drawing on the lexicon entries of either variant. (In derivational frameworks, this analysis corresponds directly to the loss of some feature driving verb movement from a lower to a higher head position.) Apart from their differential syntactic category label, conservative and innovative possessive *have* can be regarded as essentially synonymous.

Putting all of these pieces of information together, one arrives at the lexicon entries for possessive *have* shown in (125) below.

(125) a. *Conservative variant:* have *is of category I*

$$\begin{array}{ll} \textit{have}\ \texttt{I} & (\uparrow \texttt{PRED}) = \texttt{'HAVE <}(\uparrow \texttt{SUBJ})\ (\uparrow \texttt{OBJ}_\theta)\texttt{>'} \\ & (\uparrow \texttt{VFORM}) = \texttt{fin} \\ & (\uparrow \texttt{TENSE}) = \texttt{present} \\ & \neg\ \{\ (\uparrow \texttt{SUBJ PERS}) =_c \texttt{3} \\ & \quad\ \ (\uparrow \texttt{SUBJ NUM}) =_c \texttt{sg}\ \} \end{array}$$

b. *Innovative variant:* have *is of category V*

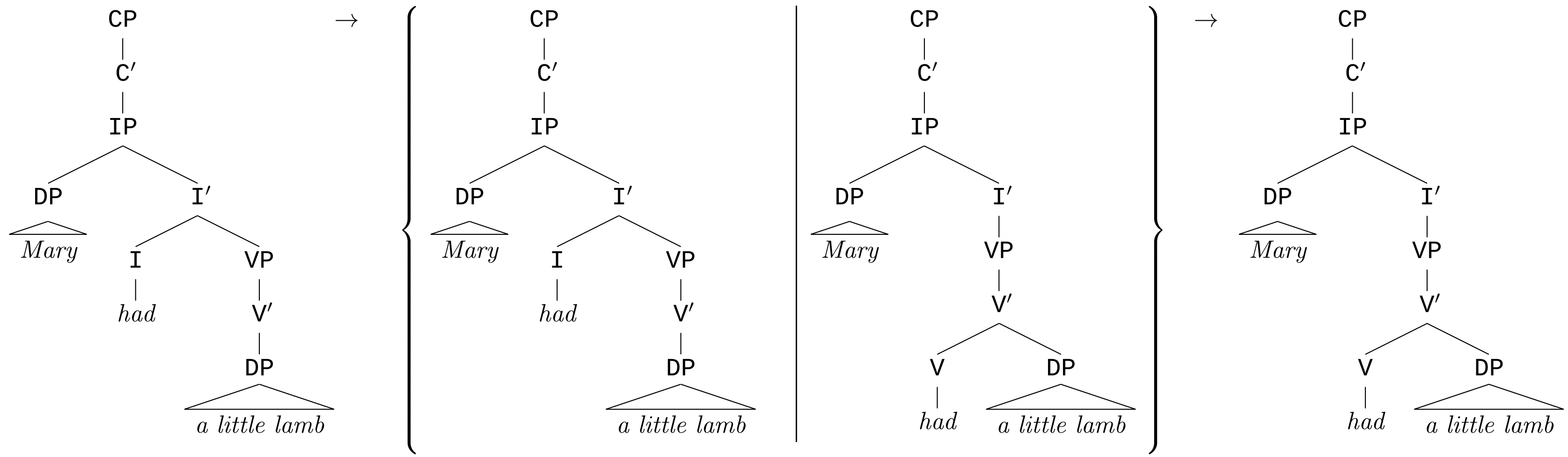
$$\begin{array}{ll} \textit{have}\ \texttt{V} & (\uparrow \texttt{PRED}) = \texttt{'HAVE <}(\uparrow \texttt{SUBJ})\ (\uparrow \texttt{OBJ}_\theta)\texttt{>'} \\ & (\uparrow \texttt{VFORM}) = \texttt{fin} \\ & (\uparrow \texttt{TENSE}) = \texttt{present} \\ & \neg\ \{\ (\uparrow \texttt{SUBJ PERS}) =_c \texttt{3} \\ & \quad\ \ (\uparrow \texttt{SUBJ NUM}) =_c \texttt{sg}\ \} \end{array}$$

The lexicon entries will force conservative possessive *have* to be positioned higher in the syntactic structure than its innovative counterpart. As long as structures can be generated with both the conservative and the innovative lexicon entries, the positional difference will directly result in variable word order patterns in accordance with the formal grammar model developed in the last sections. In other words, numerous surface alternations occurring with possessive *have* can be reduced to a single, minimal change in one of its lexical properties, i.e., its syntactic category membership. I should like to mention that the exposition of a simple causal relation between a syntactic category change and resulting surface constructions as well as the unification of a fair number of such constructions under one cause gives substantial explanatory power to the theory of the change in possessive *have* thus developed.

I would also like to stress that the analysis of conservative possessive *have* as a member of the same category as prototypical Present-Day English auxiliaries does not entail that conservative possessive *have* must itself be labeled an auxiliary. As far as I can see, nothing important hinges on the terminology used. Researchers who insist on one particular nomenclature can pick whichever technical term they feel most comfortable with. One might insist that conservative possessive *have* should not be called an auxiliary perhaps because it does not occur with a second verbal element in its clause or because it expresses a semantic relation between its arguments. One would also be justified in designating the word as an auxiliary given its peculiar linguistic behavior (e.g., just like auxiliaries, it can attract contracted negation; it has light verb uses; it has an idiosyncratic subcategorization frame different from ordinary transitive verbs, etc.).

By way of summary, the graphic below applies the general model of the implementation of a linguistic innovation in a group of speakers from chapter 1 to the specific case of the change in American possessive *have* under investigation here.

| Conservative period: Speakers represent only the conservative variant $\alpha$ | Transitional period: Speakers represent both the conservative variant $\alpha$ and the innovative variant $\beta$ | Innovative period: Speakers represent only the innovative variant $\beta$ |
|---|---|---|
| $\alpha \rightarrow$ | $\{\ \alpha \mid \beta\ \} \rightarrow$ | $\beta$ |
| Conservative period: Speakers represent possessive *have* only as category `I` | Transitional period: Speakers represent possessive *have* as both category `I` and `V` | Innovative period: Speakers represent possessive *have* only as category `V` |



#### 2.2.2.2 Predictions about the Change

The proposed category change in possessive *have* predicts, in conjunction with my formal grammar model, surface changes in four domains, namely in negation, inversion, adverb placement and ellipsis structures. I will now briefly review each of them.

First, the introduction of negation necessarily requires the presence of a head of category `I`, both for contracted negation, (25), as well as for free form *not*, (35). Thus, as long as possessive *have* is of category `I`, it can be directly negated, but as the word comes to be represented more commonly as an element of category `V`, it should increasingly be negated with *do*-support. The relevant variation is illustrated in (126).

(126) Negation

a. *Conservative variant: direct negation*
There, now, add the salt and pepper fixings, and the king himself **hasn't a slicker supper**.
*A Romance of the Mohawk*, 1840

b. *Innovative variant:* do-*supported negation*
The farming community of 900 people **doesn't have a single fast-food restaurant**.
*Weight loss x 2*, 2005

Secondly, subject-auxiliary inversion is modeled by a phrase structure rule that re-writes the `C`-head in pre-subject position to an element of category `I` in interrogative, (42), or non-canonical declarative clauses, (61). Hence, while possessive *have* is an element of category `I`, it will be placed under `C` and thus directly invert with the subject. As the item is more often realized as an element of category `V`, it cannot be inserted under `C` anymore and *do*-supported inversion becomes the norm. This variation is exemplified for subject-auxiliary inversion in positive direct questions in (127).

(127) Inversion

a. *Conservative variant: direct inversion in questions*
**Has** every person a right to keep and bear arms?
*The Revised Statutes of the State of New York*, 1843

b. *Innovative variant:* do-*support in questions*
**Do** we **have** a right to rage?
'Generations of Silence' (in: *Harper's Magazine*), 1973

Thirdly, adverbial elements frequently adjoin to `VP`, (90b), (111b). Therefore, conservative possessive *have* of category `I` will often appear in front of adverbs. Innovative possessive *have* of category `V`, on the other hand, will be positioned after adverbs. The predicted change is illustrated by the variable order of the adverb *usually* relative to possessive *have* in (128).

(128) Adverb Placement

a. *Conservative variant: post-verbal adverb*
It is within my personal knowledge that she **had usually** the charge of the house in their absence
*The History of Mary Prince, a West Indian Slave*, 1831

b. *Innovative variant: pre-verbal adverb*
These shelters **usually have** plenty of mesh in the interior canopy
'Tents' (in: *Backpacker Magazine*), 2001

Fourth, verb phrase ellipsis is regulated by phrase structural rules that truncate a syntactic tree at `VP`-nodes, (116). Consequently, possessive *have* of the conservative `I`-variety can license verb phrase ellipsis directly. Its innovative analogue of category `V`, however, loses this ability. Instead, *do*-support becomes obligatory for verb phrase ellipsis targets that are reconstructed on the basis of source clauses involving possessive *have*. The hypothesized development is exemplified in (129).

(129) Ellipsis

a. *Conservative variant: direct verb phrase ellipsis*
"You forget that I **have no money**." "I **have**, and will pay your passage."
*Hector's Inheritance*, 1885

b. *Innovative variant:* do-*supported verb phrase ellipsis*
"Renee says you **have a foolproof method** now." "I **do**. I have a caterer."
*Dean's December*, 1982

Now, the grammar model allows the deduction of a much more dramatic prediction. That is to say, the Constant Rate Hypothesis applies straightforwardly to the hypothesized change in the realization of possessive *have*. The reasoning is as follows. If all of the four above surface developments do indeed result from the same hypothesized minimal change in possessive *have*'s syntactic category, then an increase or decrease in the base probability of innovative possessive *have* should be reflected to the same extent wherever it is used. As a consequence, the rate of change from conservative possessive *have* of category `I` to innovative possessive *have* of category `V` should be identical in negation, inversion, adverb and verb phrase ellipsis contexts.

(130) *The Constant Rate Hypothesis for Possessive* have
The diachronic change in the realization of late Modern American English possessive *have* should proceed at the same rate of change in all contexts in which the item is used, i.e., in negation, inversion, adverb and verb phrase ellipsis structures.

This is the principal hypothesis to be investigated in this chapter.

#### 2.2.2.3 Relevance of the Change

The change in the realization of possessive *have* constitutes a nearly ideal test case for the investigation of the Constant Rate Hypothesis. The reasons for this concern both theoretical grammar assumptions as well as empirical contingencies of the change.

The advantages of the change from the perspective of my theoretical model are as follows. (i) The contexts used for the investigation of the Constant Rate Hypothesis are not just arbitrary divisions but can be deduced on principled grounds from a theory of Modern English clause structure. Specifically, the relevant contexts can be traced back to a fundamental difference between a higher verb position typical of Present-Day English auxiliaries and a lower verb position typical of Present-Day English main verbs. This contrasts with other test studies of the Constant Rate Hypothesis that use less soundly deduced contexts, such as a semantic distinction between temporary vs. permanent possession in Noble's 1985 study on possessive *have* in British English (cited in Kroch 1989). The outcome of the following test of the Constant Rate Hypothesis should thus be more compelling. (ii) Moreover, the model of Modern English clause structure laid out earlier is relatively uncontroversial. That is not to say that there are no more open questions regarding analyses of specific observations and their most parsimonious mathematical models, but the general tenets are well understood and standardized. As a consequence, most linguists (at least those who admit to the existence of some kind of mental grammar producing and activating linguistic structures at all) should in fact acknowledge that the various surface contexts do indeed diagnose one single underlying change and consequentially that the logic of the Constant Rate Hypothesis applies. In this respect, the test case is superior to many other linguistic changes for which it is more doubtful whether different diachronic surface variations can be related to one latent cause. For example, the general rise of *do*-support during the Early Modern English period presents a less favorable test case of the Constant Rate Hypothesis since its proper analysis is still being debated and since the change does in fact appear to involve not one cause but several distinct steps (Ecay 2015). (iii) The two competing variants differ only in one small aspect, namely their syntactic category (or in derivational frameworks, the presence or absence of some feature triggering verb movement). The change in possessive *have* thus accords much better with the prerequisites of the Constant Rate Hypothesis than would a change that does not involve near-synonymous variants, that involves more than two variants, or that does not have any competing forms to begin with.

The advantages of the change in terms of its empirical properties can be summarized in the following way. (i) The change in possessive *have* is a relatively recent development. As a consequence, there are massive amounts of relevant, well-datable textual records available and the resultant data sets allow for a high degree of statistical accuracy. In contrast, many earlier, medieval syntactic changes are not as well attested and cannot be measured with high precision. (ii) There do not seem to be any strong prescriptive pressures against the change (see Andor 2013 for a summary of attitudes towards possessive *have* in grammar handbooks). Thus, one can expect the change to be more or less adequately reflected in the textual record. Changes with strong prescriptive pressures, on the other hand, tend to retain the more prestigious variant in written documents. (iii) It is easy to identify the variants of the dependent linguistic variable. The classification of the corpus data as instances of the innovative or conservative level is thus unproblematic. There exist some changes for which the innovative, often re-analyzed variant, greatly resembles its conservative antecedent and can thus not always be unequivocally classified. (iv) The change only affects one lexical item, i.e., the verb *have* in its possessive use. One can thus ignore random effects resulting from the use of different test items.

This concludes the presentation of my theory of the change in possessive *have*. In the next parts, I will test the predictions that were shown to follow from it.

## 2.3 Measuring the Change in Possessive *have*

In the second, empirical part of this chapter, I shall conduct a large-scale corpus study to measure the rise of innovative possessive *have*. The data thus obtained can subsequently be used to test the Constant Rate Hypothesis. I will first describe the database that served as a basis for the data collection and the principles according to which it was queried. I will then consecutively summarize the possessive *have* data for the contexts of negation, inversion, adverb placement and verb phrase ellipsis. All statistical analyses were carried out in the free statistics software *R*. Specifically, logistic regression models were fitted using the functions `glm` and `lme4`'s `glmer`.

### 2.3.1 The Corpus

All data in this study was taken from the *Corpus of Historical American English* (COHA) (Davies 2010). It consists of c. 385 million words in c. 115,000 electronic files containing relatively formal, high-register, largely standard American English written texts. Every file is indexed for a relatively reliable year spanning the time period from 1810-2009. The year variable is usually based on the year of the original publication but I found a few texts for which a year of a later edition seems to have been used. Furthermore, every text file is associated with a name and one of four generic genre variables, namely news, magazine, non-fiction or fiction (for more information on the corpus structure, see Davies 2012).

The corpus can be searched conveniently through an online query interface. The queries are string-based. For example, a query such as `has he the` will return all instances of this string in the corpus. The corpus also includes part of speech tags of unknown accuracy but is not parsed. For instance, `[mc*]` is the tag used for cardinal numbers so that the search query `do not have [mc*]` will match all instances of the initial string followed by an element annotated as a cardinal number, such as *do not have one*, *do not have five*, *do not have 370* etc. A minus sign in front of a part of speech tag will match any word that is not annotated as an element of this word class. It is possible to force the presence of a single second string within a context of up to 9 words to the left or the right of the search query. Sets of several words can be searched for by conjoining them with a pipe, |. Finally, the corpus includes a wild-card option, *, which can be used to match any kind of word. All results can be displayed as Key Word in Context (KWIC) concordance lines from which the raw data can be copied and additional context can be retrieved if necessary.

The COHA is a suitable database for the purposes of this chapter on account of its substantial size, the time period it covers and the predictor variables one can retrieve on its basis. However, it is not without blemish. (i) The biggest problem I have encountered during my work with the corpus concerns duplicate results. Duplication of data usually stems from the inclusion of multiple editions of the same text and leads to inconsistent indexing of the year variable. As a consequence, a token may be annotated for a time period which it should not actually be regarded as representative of. For example, the search query `We do n't have a` will return, among many others, the following three identical hits (contexts slightly shortened).

| | | | |
|---|---|---|---|
| (131) | 1991 NYT | NEWS | clothesline is out there, too. # We do n't have a front walk to get to the |
| | 2004 Ms | MAG | clothesline is out there, too. # We do n't have a front walk to get to the |
| | 2009 MeAmpEmma | FIC | clothesline is out there, too. # We do n't have a front walk to get to the |

They are associated with three different years, 1991, 2004 and 2009, and three different genres, news, magazine and fiction, respectively. Other duplicates I have discovered can cover even wider time spans. It is also possible for a conservative variant to be cited in a more recent text file, such as the hit retrieved by the search `had you not [r*] a` in the following example.

(132) 2004 CriticalMassHow NF But pray tell me, madame , had you not formerly a more sublime idea of the universe? " - Bernard Fontenelle (1686)

Here, a conservative possessive *have* question is annotated for a recent year, 2004. The relevant token is, however, cited from a much earlier, 17$^{th}$ century text. In general, I removed duplicates whenever I found them in my data and kept only the earliest occurrence of a data point. (ii) Other problems are brought about by the nature of the search interface. For example, there is a length limit of the search query; the interface may stall if there are too many results; sometimes results are not displayed at all, etc. These are mostly minor inconveniences but may result in the omission of some data. (iii) The corpus does not include speaker and dialect variables. As a consequence it is not unlikely that certain texts were authored by British or non-native speakers of English. (iv) Finally, there are some cases of problems with the optical character recognition so that some text is not legible. All of these problems may introduce some noise into my data sets.

Henceforth, I will cite all examples from the COHA by file name and year.

### 2.3.2 Principles of Search Scripting

I followed four general principles in constructing the search queries for the data collection. They are the following.

(i) The principle of guided searches: Every search query targeted one and only one specific construction. For example, the search `has not the` specifically targeted negative declaratives, like *She has not the slightest idea*, but not questions, such as *Has not the citizen a right.* Results were removed if they did not match the targeted structure even if the retrieved tokens were relevant for another context.

(ii) The principle of symmetric searches: Every pattern was given an equal chance to be found with the innovative and the conservative form. Hence, every search was conducted twice with only a minimal difference between the conservative and innovative forms. For example, I retrieved negative declarative possessive *have* structures with a search query such as `has not the` for conservative direct negation, and hence I also conducted a minimally different search such as `does not have the` for innovative *do*-supported negation.

(iii) The principle of nearly perfect precision: Precision errors are introduced by tokens that were included in a data set but should not have been. My goal was to construct data sets with a precision that approaches 100%. In order to assure such accuracy, the result of every search query was carefully evaluated and false positives were manually deleted.

(iv) The principle of adequate recall. Recall errors are caused by tokens that were not included in a data set but should have been. I attempted to collect very large data sets of possessive *have* structures for every context. Every context was therefore targeted by many different search queries to create large data sets of at least 1,000 examples each. It was in fact often possible to retrieve substantially more examples than that. However, it is impossible to avoid recall errors entirely with the infrastructure provided by the corpus. Instead, there is a trade-off between improving recall and the time and technology investment needed to accomplish it. Thus, I had to finalize the data collection for each context at some point based on a relatively subjective feeling of sufficiency.

### 2.3.3 Negation

I will now describe how I collected the possessive *have* data for the negation context. I will then analyze the resultant data set.

#### 2.3.3.1 Data Collection

The search queries for the negation data set were assembled as follows. The dependent, linguistic variable had two variants. The conservative variant consisted of a finite form of *have* immediately followed by negation, '*have not.*' The innovative variant consisted of a finite form of *do* immediately followed by negation and the word *have*, '*do not have.*'

The specific form of the search queries varied according to four independent variables. Firstly, an adverb may or may not have been present in the negation structure. The part of speech tag for adverb in the COHA is `[r*]`. The adverb variable thus had two levels - either `[r*]` was included after the finite verb or not. The lexeme of the adverb was manually added as a variable to the data set as well.

Secondly, negation can be contracted to an auxiliary or occur as free form *not* (see section 2.2.1.3). The morphological form of the negator thus had two variants. Negation could be realized without contraction and was searched for as `not` in this case. Alternatively, negation could be contracted and was then retrieved by searching for `n't`.

Third, the finite verb can appear with different inflections. The variable could take on three levels, 'Present' for the searches involving the finite forms `have` and `do`, 'Present 3$^{rd}$ singular' for `has` and `does` and 'Past' for `had` and `did`.

Finally, the presence of an object element after *have* was required to increase the likelihood that *have* would indeed be used in its possessive function. There were ten different options for this object element: *the*, *a*, *an*, a strong quantifier that can not occur in an existential construction, a weak quantifier that could occur in an existential construction, a cardinal numeral, a demonstrative, a singular noun, a plural noun and a possessive adjective. Table 2.2 summarizes this variable by listing the names of the object elements in the order just listed, a description of the object that was targeted by this element, and their COHA search terms.

| Variant name | Targeted object | Search term |
|---|---|---|
| The | definite description | `the` |
| A | indefinite nominal | `a` |
| An | indefinite nominal (before vowel) | `an` |
| QuantStrong | quantified expression that cannot occur in existential construction | `every|each|all|most|least|both|neither` |
| QuantWeak | quantified expression that can occur in existential construction | `no|none|few|little|fewer|less|some|several|various|any|enough|more|many|much|plenty|lots` |
| Num | cardinal number nominal | `[mc*]` |
| Dem | demonstrative nominal | `this|these|that|those` |
| BareSg | bare singular | `[*nn1*]` |
| BarePl | bare plural | `[*nn2*]` |
| Poss | possessive adjective nominal | `[app*]` |

Table 2.2: The levels of the 'object type' variable in the negation data set

The data is thus based on a total of 2 (dependent variable) * 2 (adverb) * 2 (negation) * 3 (inflection) * 10 (object type) = 240 search queries.[5] For example, the search query `had n't every|each|all|most|least|both|neither` targeted the conservative variant of possessive *have* without an adverb, with contracted negation, past inflection and a strongly quantified object; the search query `does not [r*] have the` targeted the innovative variant of possessive *have* with an adverb, full form negation, 3rd person singular inflection and a definite object, etc.

Once all the search queries were run, I manually corrected precision errors. The following were the most common kinds of false hits that I removed from the data. The targeted object often turned out to be a completely different element, such as a badly tagged past participle, (133a), (133b), a measure phrase, (133c), or other kind of adjunct, (133d). In the following examples, I underline the string erroneously matched by a specific search query.

(133) a. if he <u>has not lock'd</u> the door
GordianKnot 1810
b. do think of it – <u>had n't a</u> been for him I should a slipped cable
TennesseanANovel 1827
c. you <u>have not much</u> longer to bear with my humours
YankeyInEngland 1815
d. if Lloyd <u>had not the</u> next day made it the text upon which he preached
New-EnglandTale 1822

The targeted object was also frequently the subject in inversion contexts such as questions, (134a), or verb-first conditionals, (134b).

(134) a. <u>Has not his</u> sire With impious step invaded all our temples?
AlexisCzarewitz 1812
b. and he might have fallen, <u>had not the</u> bishop stretched out his hand
JourneyInOther 1894

I also deleted all non-finite forms of possessive *have*, as in the following examples.

(135) a. a large part of our assemblies seem to <u>have not the</u> least notion
NorthAmRev 1822
b. and I have <u>had not a</u> few of their mortal foes under my roof
RachelDyerANorth 1828

Finally, I removed instances of small clauses, in which the targeted object appeared to be an argument of a following predicate. There was some subjectivity in determining whether elements should count as offending predicates or not. Examples are shown below.

[5] The exact search queries can be viewed on the Dissertation DVD under `Chapter 2 - Possessive Have/03 Negation Context/Negation Queries.xlsx`.

(136) a. The tax for schools, being collected with the county tax, had not the odium attached to it of a fine
NorthAmRev 1823

b. Miss Atherton had not this temptation to contend with.
ThirtyYearsAgo1836

c. bees do not have the wings hooked together
Atlantic 1925

The manual correction of precision errors proved quite a tedious and strenuous task. It is therefore possible that several false hits were not spotted and deleted. Nevertheless, the vast majority of examples in the negation data set should correspond to the definition of possessive *have* used here.

The data collection resulted in a structured, highly accurate data set with a total of 33,074 examples.[6] The two examples below illustrate the kind of sentences included in the data set with some near minimal pairs.

(137) a. *conservative variant:* 'have not'
But we **have not space** here to recount what is known about the linguistic picture of pre-Indo-European
ScienceLanguage 1962

b. *innovative variant:* 'do not have'
... because some languages **do not have words**, and because most words have more than one semantic element
ScienceLanguage 1962

(138) a. *conservative variant:* 'have not'
... whether the General Government have or **have not a** right to lay out roads and canals
NorthAmRev 1831

b. *innovative variant:* 'do not have'
... in order to determine if they do or **do not have** a conscious or unconscious prejudice.
Time 1964

#### 2.3.3.2 Results

I will explain the statistical evaluation of the negation data set in relatively great detail. The line of argument can be applied directly to the other contexts.

I fitted a logistic regression model to the data. It regresses the linguistic variable as the log-odds of *'do not have'* vs. *'have not'* against time to trace the general development of possessive *have* in negative declarative sentences. The resultant model is shown below.

```
formula = Have ~ Year, family = binomial, data = negation
```

| | Estimate | Std.Error | z-value | *p* |
|---|---|---|---|---|
| Intercept | -80.09 | 0.8240 | -97.20 | $<0.001^{***}$ |
| Year | 0.04137 | 0.0004246 | 97.42 | $<0.001^{***}$ |

```
Null deviance:        45792 on 33073 degrees of freedom
Residual deviance:    27286 on 33072 degrees of freedom
AIC: 27290
```

Table 2.3: Logistic Regression Model 1: possessive *have* against time in negative declaratives

The model confirms the assumption that *do*-support in negative declaratives becomes the norm during the late Modern period in American English. The effect of time on the variation between *'have not* and *'do not have'* is measured by the coefficient for the Year variable. Here, the log-odds of the innovative variant *'do not have'* increase by 0.04137 for every one year increase in time. At this rate of change, it would take 223 years for the innovative form to rise from 1% to 99% of use. The standard error of the Year coefficient is 0.0004246, which results in a 95%-confidence interval of [0.0405 – 0.0422]. The Year variable is a highly significant predictor for the realization of the linguistic variable in this model.

The data and the above model are graphically illustrated in figure 2.3 below.

[6]The negation data set is stored on the Dissertation DVD under `Chapter 2 - Possessive Have/03 Negation Context/Negation Data.xlsx`.

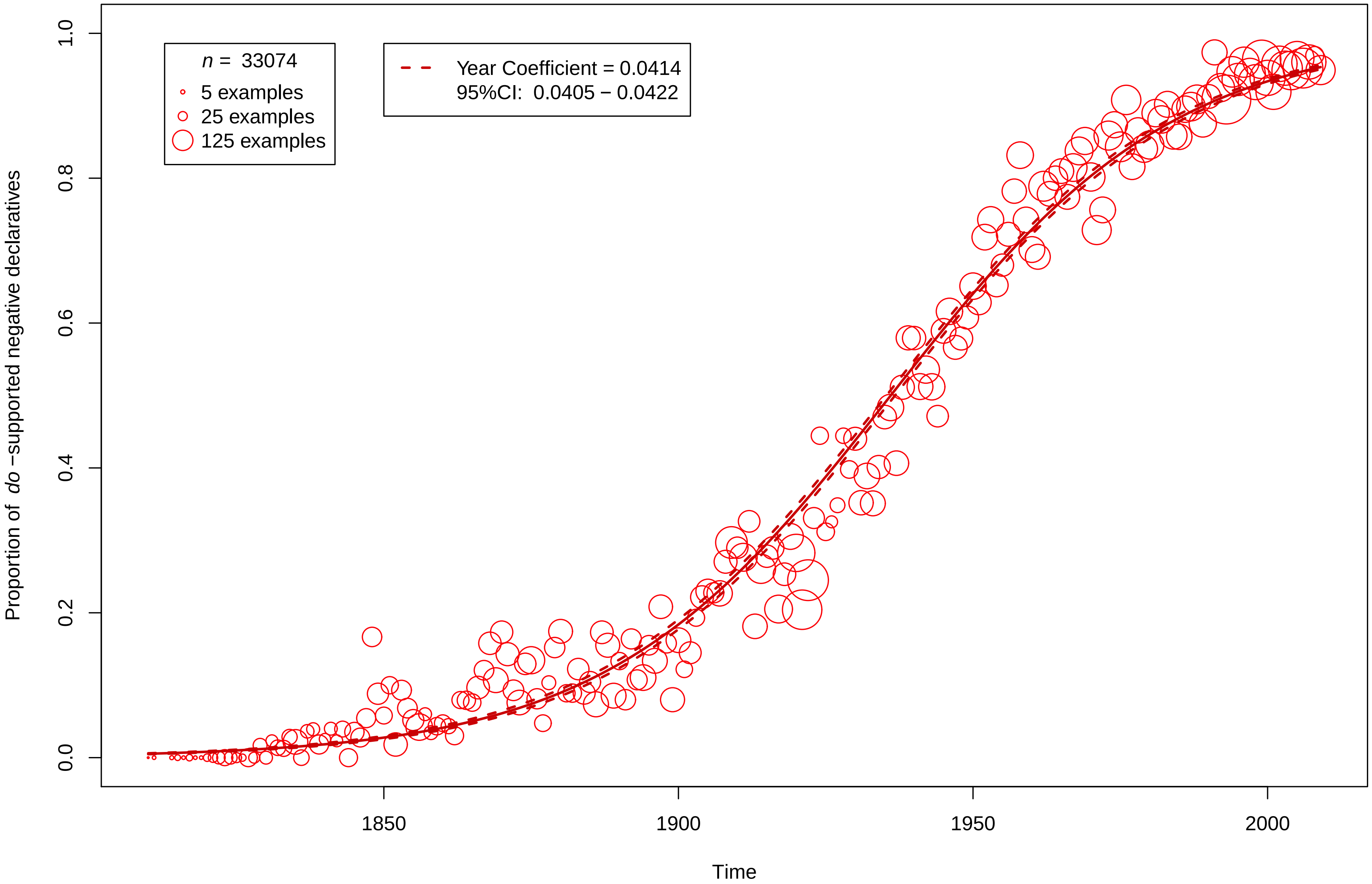



Figure 2.3: Development of possessive *have* in the negation context

The graph plots the proportion of innovative, *do*-supported negative possessive *have* declaratives against time. There is enough material in the data set for 198 data points, i.e., a reference to every single year from 1810-2009 except for the years 1812 and 1813. The size of a data point is proportional to the overall number of examples for that year such that a default point character corresponds to 25 examples. The logistic regression fit is represented as a line drawn through the point cloud. Its slope, hence the effect size of time or, alternatively, the speed of the change, is determined by the Year coefficient. The narrow dashes indicate the 95%-confidence bands for the fitted regression line.

The graph shows a clearly visible s-shaped curve for the spread of the innovative variant *'do not have.'* It has often been reasoned that linguistic changes should follow an s-shaped, logistic growth pattern, spreading slowly at first, then speeding up, and finally slowing down again as the innovative form reaches full penetration in the population and hence reaches its upper limit (e.g., Weinreich et al. 1968, Altmann 1983, Kroch 1989). In practice, however, researchers have often merely superimposed an s-shaped pattern on, rather than actually observed it in, their data. The present data set, in contrast, incontrovertibly features a distinctly s-shaped curve. The clarity of this pattern is due to the large size of the data set, which includes more than 33,000 examples, and which may in fact be on one of the largest structured data sets of a syntactic change ever collected. The observation of an s-shaped curve in my data thus supports the commonly held assumption of logistic growth in linguistic changes.

Given its simplicity, the model performs in a very satisfactory way. The model fits significantly better than a model with just an intercept, i.e., a null model (Likelihood Ratio Test, $\chi^2$=18505.84, df=1, $p$<0.001). The model fit is good (Nagelkerke Pseudo-$R^2$=0.572). The model's classification accuracy is high (C-index=0.894, classification accuracy = 82.6% vs. baseline: 52.1%).

However, an analysis of the residual plot, which displays for every data point the divergence between the observed and the estimated values of the dependent variable, suggests that there may be important generalizations in the data that are not captured by the model. The residual plot is shown in figure 2.4 below, where a positive residual indicates a greater than expected proportion of *'do not have'*, a negative residual a greater than expected proportion of *'have not.'*

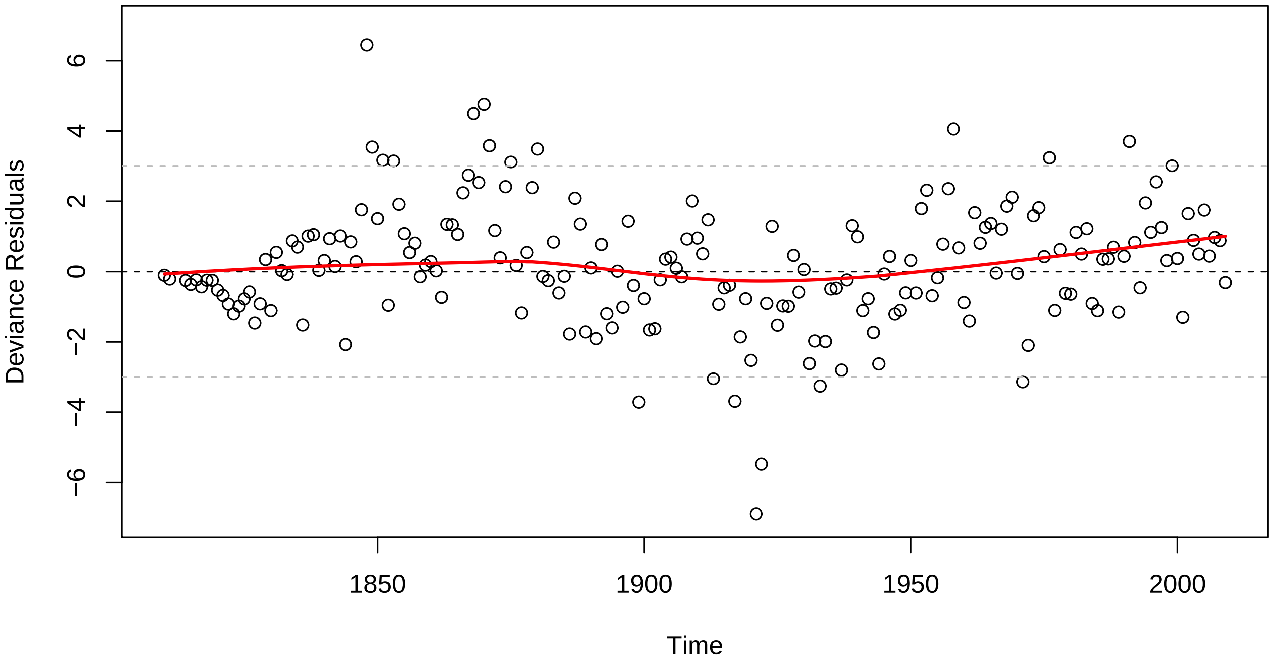



Figure 2.4: Residual plot of logistic regression model 1

If the fitted logistic regression model was appropriate, most of the residuals should fall within a range of ± 3, indicated by the gray dotted lines. Since there are many observations of the dependent variable for every year, the distribution of the residuals should approximate to normality. Consequently, the residuals should form an unstructured, horizontal band around the estimated value of the dependent variable, where the residuals are 0. Here, about 10% of all data points show unexpectedly extreme residuals. Specifically, there are many outlier years with a greater than predicted chance of *do*-supported negative declaratives in the early stage of the change, between c. 1845 and 1880. At the beginning of the 20th century, there are a few outliers to the downside, with more instances of *'have not'* than expected. Finally, a few dispersed outlier years with a greater than expected probability of *'do not have'* appear again towards the later stages of the change. The residuals are clearly not normally distributed (Shapiro-Wilk normality test, W=0.97442, $p$=0.0011). Instead of an unstructured, horizontal band, the residuals form a slightly curved, s-shaped pattern, as highlighted by the red smoothing line. This suggests that another factor exerts a decisive influence on the variation, leading to a non-linear relationship between time and the realization of possessive *have*.

I believe that the unexpected residual pattern can be explained in the following way. The COHA predominantly includes texts of a high, prestigious, standardized register of American English, which is typically used by the upper, educated social classes in formal contexts. Within that register, however, non-standard forms may sometimes be integrated, quoted or parodied. Vernacular forms are commonly found, in particular, in direct speeches of fictional texts, where they reflect not the grammar of the authors, but a caricature of one of their characters' manner of speaking. If such register fluctuations are frequent enough, they may account for the outlier years.

The innovative variant of possessive *have* seems to have functioned as a low-prestige, colloquial form outside of standard American English in the early 19th century. The form was recorded in direct speech of fictional texts from that time as a means to portray the social standing of lower-class literary characters. This intrusion of innovative possessive *have* from a low-prestige variety into high-register fictional texts from the early stages of the transitional period could then substantially elevate the overall frequency of that form in those documents, and hence distort the estimated probability with which the form would be expected to occur in textual records from that time. This in turn may result in the observed early outlier years with a greater than expected chance of finding the innovative form and upward curving residuals.

The examples in (139) present a small selection of a large number of sentences that could be cited as anecdotal evidence in support of this explanation. They illustrate innovative possessive *have* structures in low-prestige, vernacular direct speeches from fictional texts.

(139) a. Dea'! I don't see why they **don't have an elevata**
OakOpenings 1848

b. ef we **don't have a supply o' water**, we're likely to perish
EllaBarnwell 1853

c. I've seed picters of this place before, but I **didn't have no idee** it was so handsum
MajorJonessSketches 1848

d. They **don't never have no doubt** o' your meaning, Tom.
UncleTomsCabin 1852

e. dey was little boys, and I **didn't have much trouble** in gettin' 'em here
LiberiaMrPeytons 1853

f. they **don't have no dinners**, but keeps a dinin' all day off fat pigs
RedburnHisFirst 1849

g. "Wal, wal," said the captain, "I **didn't have much hopes**; it's jest as I feared."
LostInFog 1870

The above discourses betray their vernacular origins through many indicators. There are colloquial exclamations (*Dea'*, *Wal, wal*), unusual indications of reductions and contractions (*o'*, *dinin'*, *ef*, *'em*), a regularized strong past tense (*seed*), non-standard spellings of words (*elevata*, *idee*, *handsum*, *dey*), multiple negations (*didn't have no*, *don't never have no*) etc. Further, they do not reflect the author's way of speaking but rather mirror the vernacular of fictional characters. The speaker in (139a) is a woman of common origin whose "accent and [...] diction gave her away for a middle-class New England person of village birth" (from c. two pages before the above citation). Sentence (139b) is spoken by a fictional, old army veteran from an isolated fort at the early American Frontier near Lake Erie. The third example, (139c), is uttered by William T. Thompson's literary character Major Joseph Jones, a whimsical middle-class plantation owner from Pineville, Georgia. Sentence (139d) is taken from the famous book *Uncle Tom's Cabin* and is spoken by Marks, a (presumably Southern) slave hunter, hired to track down fugitive slave Eliza. Sentence (139e) is uttered by an old, freed slave, who has settled in Canada. The speaker of (139f) is a whaler from Nantucket Island, Massachusetts, who hates civilized society, and the last sentence, (139g), is evidently exclaimed by a clichéd sea-captain.

The fact that *do*-supported negative declarative possessive *have* sentences often feature in the portrayal of lower classes, uneducated speakers or marginal groups may indicate that the change spreads from, and maybe even originates in, speakers of relatively low socio-economic status. This scenario is also compatible with a change 'from below' in the technical sense, where a new form typically (though not necessarily) enters the main variety through a vernacular, carries no overt prestige, remains below the level of speakers' conscious awareness, and is advanced primarily by principles of language acquisition by young children (e.g., Labov 2007). Stronger claims about the socio-linguistic covariates of the change and the mechanism of its social implementation seem to me unwarranted on the basis of the data available here.

By the first half of the 20th century, innovative possessive *have* may either have spread so much that it lost its value as a literary device for the depiction of the lower socio-economic classes, or fictional writing may have undergone a shift towards a less dramatic style with fewer imitative direct speech passages. Consequently, the records now include fewer examples of innovative possessive *have* than would be predicted by a regression based on the earlier, elevated incidence of *do*-support. This may explain the outlier years with a lower than expected chance of finding the innovative form and the downward-sloping curvature in the residuals observed at that time.

Finally, in the later stages of the transitional period, a reverse register effect might obtain. The conservative variant of possessive *have* could now be associated with conspicuously archaic, quaint, formal or British connotations, and be exploited for literary purposes accordingly. That is to say, authors could either retain conservative possessive *have* structures to convey a certain style, say a formal rhetoric, a pastoral literary mode, an indexing mechanism for aged speakers etc., or put such forms in the mouths of their literary characters directly in an effort to make them appear old-fashioned. This would lower the observed proportions of innovative forms towards the second half of the change. Some years for which the COHA happens to include few such antiquating fictional texts might then register with a higher, untainted proportion of innovative possessive *have.* This might explain the unevenly scattered late outliers with a higher than expected chance of finding the innovative form and upward sloping residuals.

Anecdotal evidence of this register effect is harder to ascertain because there are few unequivocally archaic features which conservative possessive *have* could co-occur with. The effect can nevertheless be supported by some fairly clear instances of conservative possessive *have* structures in recent pastiche texts or historical fiction.

(140) a. Two is something else entirely. I fear we **have not a moment** to lose.
DustShadowAccount 2009

b. "To be frank, Lady Phaedra, Papa **hadn't any taste**."
TemptedAllNight 2009

c. My mother **has not the will** – but you know her trials. I must wed without delay
RoguesSalute 2007

Example (140a) comes from a modern Sherlock Holmes novel by Lyndsay Faye, who consciously imitates Arthur Conan Doyle's historical style and even enjoys his heirs' official endorsement. Sentence (140b) is taken from Liz Carlyle's historical novel *Tempted All Night*, which is set during the regency period in England of the first decades of the 19th century. The last example, (140c), occurs in the historical romance *Rogue's Salute* under a chapter heading entitled "New Orleans, Louisiana, January 1842." It stands to reason that the conservative possessive *have* forms in these examples are retained so that the literary characters would accord with the diction imagined as typical of the time portrayed.

If the irregular residual pattern in figure 2.4 is in fact due to the effect of register fluctuations as just described, one would expect the following predictions to be true. Two logistic regression models fitted separately for non-fictional vs. fictional texts should reveal two independent developments and residual patterns. (i) In the first stages of the change, non-fiction should show very few instances of innovative possessive *have* but fictional writings should occur with a higher proportion of such structures due to recordings of vernacular forms. In the later stages of the change, non-fiction should involve very many instances of innovative possessive *have* whereas fictional texts should have a lower proportion of such structures because of the stylistic retention of archaisms. As a consequence of this imbalance, the time curve should be steeper for non-fictional texts and flatter for fictional texts and hence the Year coefficient should be greater for the former than for the latter genre. (ii) Outlier years should largely disappear in the non-fiction data subset but persist in the fictional texts. (iii) Residuals should form a structureless, horizontal band in non-fictional texts but still show a slightly s-shaped curvature in fictional writing. The residuals should be distributed approximately normally. Here and in the following sections I operationalize non-fiction and fiction by means of genre variables provided by the COHA. Non-fiction encompasses texts that are annotated as MAG, NEWS or NF whereas fiction is regarded as the sum of texts that are annotated as FIC.

Figure 2.5 below graphically presents logistic regression models and residual plots for non-fictional and fictional writings in the left hand and right hand columns respectively.

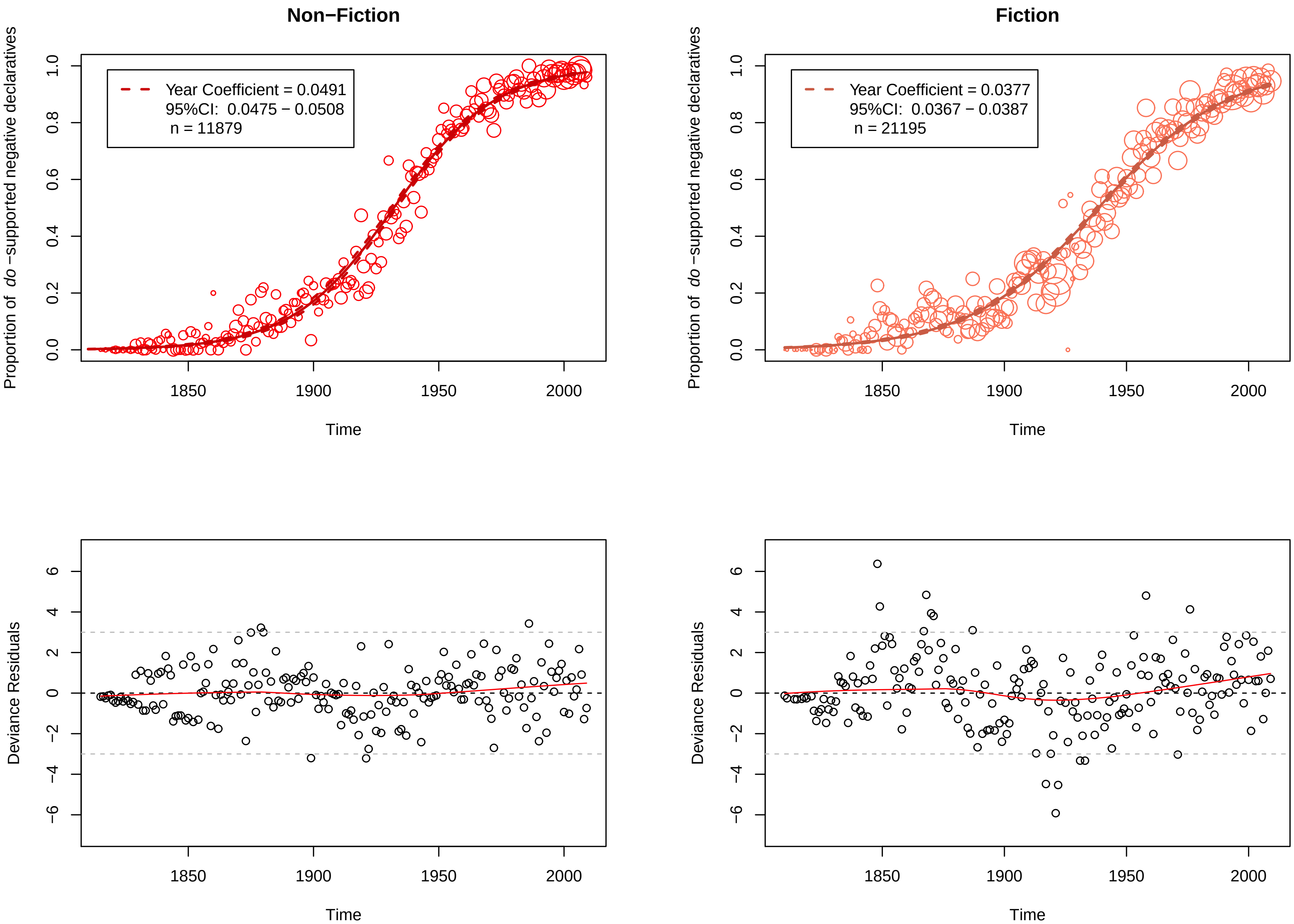



Figure 2.5: Development of negative declarative possessive *have* and residual plots for fictional vs. non-fictional texts

The results of investigating the change separately for different genre categories are in accordance with the above predictions. (i) The time curve is steeper in non-fictional texts than fictional writings. Specifically, the Year coefficient is greater in non-fiction texts at 0.0491 than in fiction texts at only 0.0377. An innovative form would rise from 1% to 99% of use in 187 years at the former rate, but in 244 years at the latter rate. (ii) Outliers do not appear frequently in non-fiction. Only 2% of all data points have a residual value more extreme than $\pm$ 3 in that category. In contrast, outlier years are still found as before in fictional texts. About 8% of all data points have unexpectedly extreme residual values of more than $\pm$ 3 in that genre. (iii) The residual plot for non-fictional writing shows a horizontal band. Its smoothing line does not reveal an obvious trend in the residuals. The residual plot for fictional texts, on the other hand, still displays structure in the residual data points and its smoothing line has a slight s-shape. The residuals in both contexts now have a much greater probability of being drawn from a normal distribution (Shapiro-Wilk normality test, non-fiction: $W= 0.99256$, $p=0.4251$, fiction: $W=0.98839$, $p=0.1108$). The confirmation of these predictions constitutes evidence for the effect of register fluctuations outlined above as a plausible explanation for the unexpected residual pattern.

The question then becomes how one should statistically model the effect of fictionalized imitations of vernacular or archaic uses of possessive *have*. The most obvious answer lies in the inclusion of the difference between fictional and non-fictional texts as a categorical variable in the regression model. One must then decide if the genre difference should be regarded as a fixed or as a random effect. While it is difficult to pinpoint precisely the difference between fixed and random factors (see e.g., Gelman 2005: 20, for 5 different definitions of 'random effect'), the distinction can roughly be understood as follows. Fixed factors consist of levels that are identified as relevant by a researcher, and hence were specifically chosen for a study. The chosen levels in a sample are usually relatively few in number, or the variable is continuous. The intent of the inclusion of the fixed factor is to extrapolate from the effect of its levels in the sample to the effect of the same levels in the system under investigation in general, or to the effect of the same levels in a new sample. Random factors, in contrast, have levels that are not actually of interest to a researcher, but were randomly sampled from a larger, potentially infinite, population. The randomly picked levels in a sample are usually very numerous. The purpose of the inclusion of the random factor is to extrapolate from the effect of its levels in the sample to the effect of the population of possible levels in the system under investigation in general, or to the effect of different levels in a new sample.

Should the genre difference between fictional and non-fictional writing be modeled as a fixed or as a random factor? It is rather difficult to conceptualize genre effects within the scope of the above definitions and both options can be encountered in the literature (e.g., Levshina 2015: ch. 13 for genre as a fixed effect; Gries 2015 for genre as a random effect). I believe that the effect under consideration in the present case is best thought of as a fixed effect. I am specifically interested in the difference between texts that consistently conform to standard American English typical of its time vs. those that mimic forms from a different variety. Indeed, fiction vs. non-fiction as used here should be considered as a mere proxy for a continuous variable that measures the degree to which texts imitate speech in fictional worlds. (Operationalization of this variable could be achieved, for instance, in terms of the frequency of the first person singular pronoun, punctuations indicating direct speech, occurrences of vernacular spellings such as *-in'* or *'em*, a bag of informative words, lexical complexity, average sentence length etc. Warner (2006) provides a similar characterization of text-internal linguistic properties that correlate with the appearance of early Modern English *do*-support with main verbs. Unfortunately, such a measurement is not feasible within the scope of this chapter.) Inversely, I do not conceive of the variable as a random sample from a larger population of possible genres. Furthermore, there are only a small number of levels of the factor, i.e., non-fiction vs. fiction. In contrast, random factors should have a larger number of levels in the data set. Finally, I would claim that the effect of the difference between authentic vs. strongly imitating styles on the realization of possessive *have* generalizes to the system of American English in general and could be replicated in other samples as well. By way of contrast, I do not wish to extrapolate from the effect of different genres found in my data set to the variation between genres in American English in general. In fact, it is statistically unrealistic to estimate how much variation one should expect between different genre categories based on only a handful of genre levels annotated in the COHA.

Accepting the above reasoning, I will thus include the genre difference between non-fiction and fiction as an additional predictor in the regression model. The genre term will allow the intercept of the regression line to vary between its two levels. However, I have also shown that the slope of the curve is markedly different between the two categories - non-fictional texts seem to develop *do*-supported negative declarative with possessive *have* faster than fictional texts. This finding can be modeled by also including a term for the interaction between the time and genre variables. The resultant logistic regression model is presented below.

```
formula = Have ~  Year + Genre + Year:Genre,
   family = binomial, data = negation
```

| | Estimate | Std.Error | z-value | p |
|---|---|---|---|---|
| Intercept | -94.85 | 1.633 | -58.20 | $<0.001^{***}$ |
| Year | 0.04910 | 0.0008421 | 58.30 | $<0.001^{***}$ |
| Genre(non-fic→fic) | 21.75 | 1.895 | 11.48 | $<0.001^{***}$ |
| Year:Genre(non-fic→fic) | -0.01139 | 0.0009768 | -11.66 | $<0.001^{***}$ |

```
Null deviance:        45792 on 33073 degrees of freedom
Residual deviance:    27036 on 33070 degrees of freedom
AIC: 27044
```

Table 2.4: Logistic Regression Model 2: *have* by time, genre and their interaction in negative declaratives

The regression output above can be interpreted as follows. The Year coefficient functions as before. In this case, the model predicts an increase in the log-odds of innovative *'do not have'* of 0.04910 for every one year advance in time. The next two coefficients are evaluated with respect to a reference value for genre, which is non-fiction in the above model notation. The 'Genre' coefficient estimates the effect of its different levels on the log-odds of the dependent variable. In this model, the log-odds of *do*-supported negative declarative possessive *have* rise by 21.75 as the genre category is changed from non-fictional to fictional texts. Finally, the interaction term between genre and time indicates the difference in the change of the log-odds of the dependent variable per unit increase in the Year variable for the different levels of the Genre variable. In this case, the model predicts that every one year increase in time will boost the log-odds of *'do not have'* by 0.01139 less if the genre category is changed from non-fiction to fiction. Note that figure 2.5 indicates in the upper left panel for non-fictional texts the above Year coefficient, 0.04910, while the slope for the graph in the upper right corner for fictional texts corresponds exactly to the difference in slope induced by the time-genre interaction, 0.04910 - 0.01139 = 0.03771. Each of the three variables is a highly significant predictor for the realization of possessive *have* in this model.

The new model performs slightly more favorably than the previous one. As before, the model fits significantly better than a null model (Likelihood Ratio Test, $\chi^2$=18756.39, df=3, $p$<0.001). The model fit is good (Nagelkerke Pseudo-$R^2$= 0.577 (before: 0.572)). The model's classification accuracy remains high (C-index=0.896 (before: 0.894), classification accuracy = 82.6% (before: 82.6%) vs. baseline: 52.1%).

Overall, the genre effect is quite weak in comparison to the time effect. This can be shown, for instance, by comparing the AICs (or BICs) of a null model, a model with only one predictor, and a model with both predictors and their interaction, all fitted to the same, unbinned data (the smaller the AIC (BIC), the better the fit of the model). Here, the null model's AIC is 45794, and reduces immensely when the time variable is added (by 18504 if added as first predictor, by 18393 if added after genre). In contrast, the AIC reduces by a much smaller amount after the addition of the genre and interaction variables (by 246 if added after time, by 357 if genre added as first predictor). The BICs show comparable results.

| | Null model (just intercept) | Model 1 + time | Model 2 + time * genre |
|---|---|---|---|
| AIC | 45794 | 27290 | 27044 |
| BIC | 45803 | 27307 | 27077 |

Table 2.5: AICs and BICs of null, time and time*genre models for the negation data

Hence, most of the variation in the realization of possessive *have* appears to be due to time, not to the difference between non-fiction and fiction. It seems plausible that the primary effect of time can by and large be observed in all genre categories. This supports the assumption that possessive *have* undergoes one simple, underlying change and that the observed difference in genre is merely the consequence of an additional layer of extra-linguistic, non-grammatical, or literary contingencies.

I would like to modify the model in one more way: I will include a random effect for individual texts in addition to the fixed effects above, which will result in a so-called mixed-effects model. Mixed-effects models have come to be used widely in areas such as sociolinguistics and psycholinguistics but seem to me to be still relatively unfamiliar in the study of language change (for some studies of linguistic change that employ mixed-effects models, see e.g., Hilpert and Gries 2010, Hilpert 2013 ch. 2, Wallenberg 2014). I will therefore spend some time to explain, in an informal way, the logic and workings of mixed-effects models and apply the relevant concepts to the change in possessive *have* at hand. I will focus my discussion on the relation between individual texts and the 'Year' variable, but the reasoning would extend equally to other effects. The account presented draws on the informal overview in Robson & Pevalin (2016) as well as some technical details from Gelman & Hill (2007).

In linguistics, it is often the case that tokens clustering together as levels of a fixed factor are actually randomly sampled from a larger population of individuals. The fixed group-level variable will then have a nesting relationship with its random individual-level variable. In the present study, the main interest lies in the effect of time, modeled as the systematic difference between years (group-level variable), which are nested in textual excerpts reflecting the grammar of a specific speaker or author sampled from a large number of documents of that time (individual-level variable). For example, the year 1899 consists of examples pooled from $n$ different texts; the year 1900 is represented by $o$ texts; the year 1901 is made up of a total of $p$ distinct texts, etc. Each text belongs to one and only one specific year. Figure 2.6 illustrates the nesting relationship between random texts grouped together into fixed years.

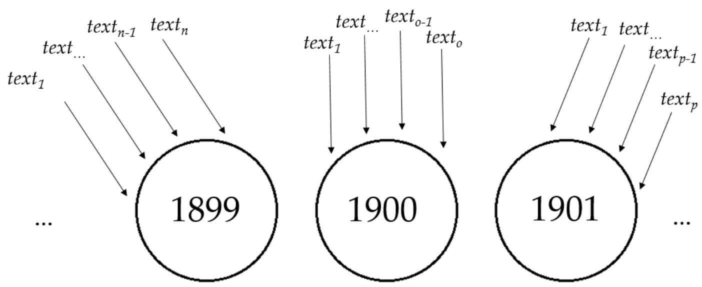



Figure 2.6: Nesting relationship between the time and text variables

The effect of the randomly sampled individuals, here texts, should not be ignored for the following reason. Regression models that exclusively incorporate fixed factors implicitly assume that there are no individual differences within the group levels of a predictor variable. In other words, there should be no correlation between the errors of any of the observations within a group. In the present case, a fixed effects model with time as the only effect would presuppose that every observation has the same chance of being realized as a *do*-supported negative possessive *have* structure within a particular year. There should be no correlation among the discrepancies between individual observations, either *'have not'* or *'do not have'*, and the expected mean of *'do not have'* for a given year. In reality, however, tokens of linguistic variables are rarely independent of each other. Instead, the errors of some observations are typically correlated among individuals randomly sampled from a larger population. Here, it seems plausible to assume that texts may systemically behave differently from one another. It is likely that there will be correlations among the errors of individual observations of possessive *have* in negative declaratives within particular texts. As a consequence, ignoring the random effect may lead to an unrealistic estimation and an exaggeration of the significance of the between-group predictors. A part of the variation in the dependent linguistic variable may be attributable not to fixed effects but to individual differences. In the case under investigation, the unwarranted pooling together of tokens from disparate texts may mis-characterize and over-estimate the effect of time. Some of the variation in the realization of possessive *have* may not be due so much to the general time effect but rather to highly influential texts, e.g., those that are large and particularly innovative or conservative for their year (for a discussion of the general usefulness of mixed-effects models in linguistics, see Johnson 2014).

In order to avoid these problems, my model should thus include a 'Text' variable to model individual differences within years. I will briefly discuss some characteristics of this 'Text' variable. The factor is composed of levels representing the names of the COHA text files. Search queries with the COHA interface automatically return the name of the text file that an observation is drawn from. Some text names occur multiple times (e.g., there are many instances of *Time* magazine). I associated every file name with its year of composition to ensure a one-to-one relation between a text and its year. It should be obvious that the variable does indeed constitute a random factor according to the definition given earlier. The different texts included in the data set should be thought of as a random sample from a much larger population of texts written between 1810 and 2009. The specific texts are not actually of any interest here. There are very many levels of the factor. For example, the data of the negation context comes from a total of 7,586 different text files. The goal of the inclusion of the 'Text' variable lies in the capability to generalize the result of the regression model from the specific texts in the data set to the population of American English texts of the transitional period in general. Further, the token contributions from the individual texts are quite heterogeneous in size. The ratio of texts to tokens in the negation data can be summarized as follows. A large number of the texts, 2,608 or 34.4%, contribute only one example of possessive *have*, and 1,269 or 16.7%, include only two such examples. Very few texts, namely only 142 or 1.9%, include 20 or more relevant examples. The text to token ratio is presented in the left panel of figure 2.7. The number of texts are also unevenly distributed across the transitional period. There are few texts for the earliest decades of the change, from c. 1810-1830, increasingly greater numbers of texts for the following decades with a dip during the 1920s, and substantially more texts from the late 1990s to the present. The right panel of figure 2.7 illustrates the temporal distribution of texts.

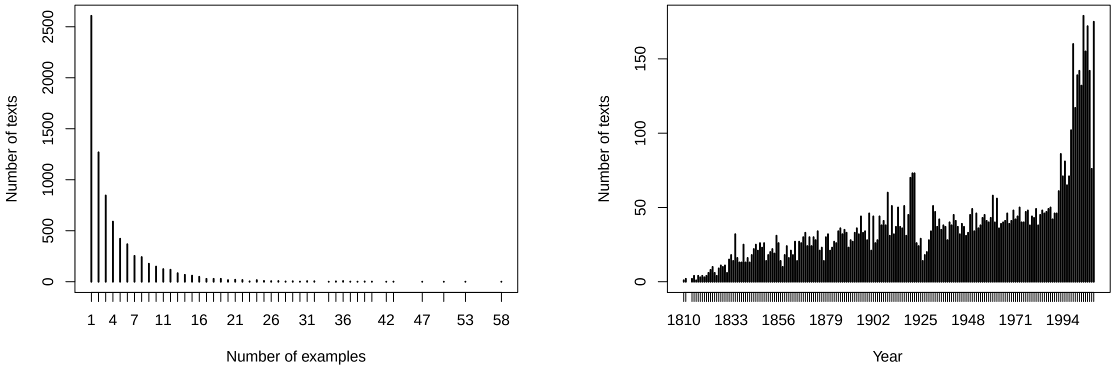



Figure 2.7: Average size and temporal distribution of texts in negation data

Finally, the ‘Text’ variable is an imperfect approximation of the underlying variation between the grammars of individual authors or speakers. On the one hand, some text levels refer to output produced by different writers, such as articles grouped together into one magazine. On the other hand, there may be several texts that were actually composed by the same author but have different labels. Nevertheless, differences between individual texts can be assumed to correlate reasonably well with differences between individual grammars.

How will the mixed-effects model factor in the individual variation in the realization of possessive *have* between texts? An informal description of the logic behind mixed-effects regression could read as follows. (i) Fixed-effects models have exactly one intercept, indicating the value of the dependent variable when all predictors are 0. Thus, the two fixed-effects models in tables 2.3 and 2.4 above involve only one intercept, which represents the overall probability of innovative possessive *have*. These models thus ignore the fact that the global mean is actually based on a large number of disparate texts. (ii) In contrast, mixed-effects models permit (among other options) random intercepts for every individual. Hence, my mixed-effects model will partition the model intercept into an average value and a text-dependent part so that every text can in principle occur with its own intercept, i.e., its own baseline probability of realizing possessive *have* as the innovative variant. Intercepts of texts that show a proportion of innovative *have* markedly different from the group mean will veer from the global intercept to a greater extent than intercepts of texts whose proportions of innovative *have* conform more closely to the group mean. This, in essence, encapsulates the variation among individual texts. Further, the intercepts of larger texts will be allowed to deviate more from the group mean than the intercepts of smaller texts. This compensates for the heterogeneity in text size. (iii) Mixed-effects models will adjust the parameter estimates of the regression equation as a consequence of the consideration of individual variation within random effects. Here, the inclusion of the ‘Text’ variable will lead to estimates of the intercept and coefficients of the between-group predictors that differ from the respective estimates of the fixed-effects models and that reflect more adequately the influence of particularly idiosyncratic texts. This will correct the potential problem of over-estimating the time predictor. (iv) In practice, the variability among the varying intercepts for every text is estimated as the parameters of a normal distribution with mean zero and a variance composed of two parts - the variability among the point estimates of the random intercepts and the uncertainty of the random intercept estimates themselves. This variance parameter is reported in the logistic regression output and can serve as an indication of the extent of variability among the randomly sampled individuals in the data.

I will now illustrate these points by way of reference to the negative possessive *have* data isolated for the year 1900. There are a total of 185 examples of possessive *have* in my negation data for that year. 30 of those examples or 16.2% occur with the innovative *do*-supported variant. The examples are drawn from a total of 46 texts. Figure 2.8 below illustrates the effect of individual texts on the variation in possessive *have* in 1900.

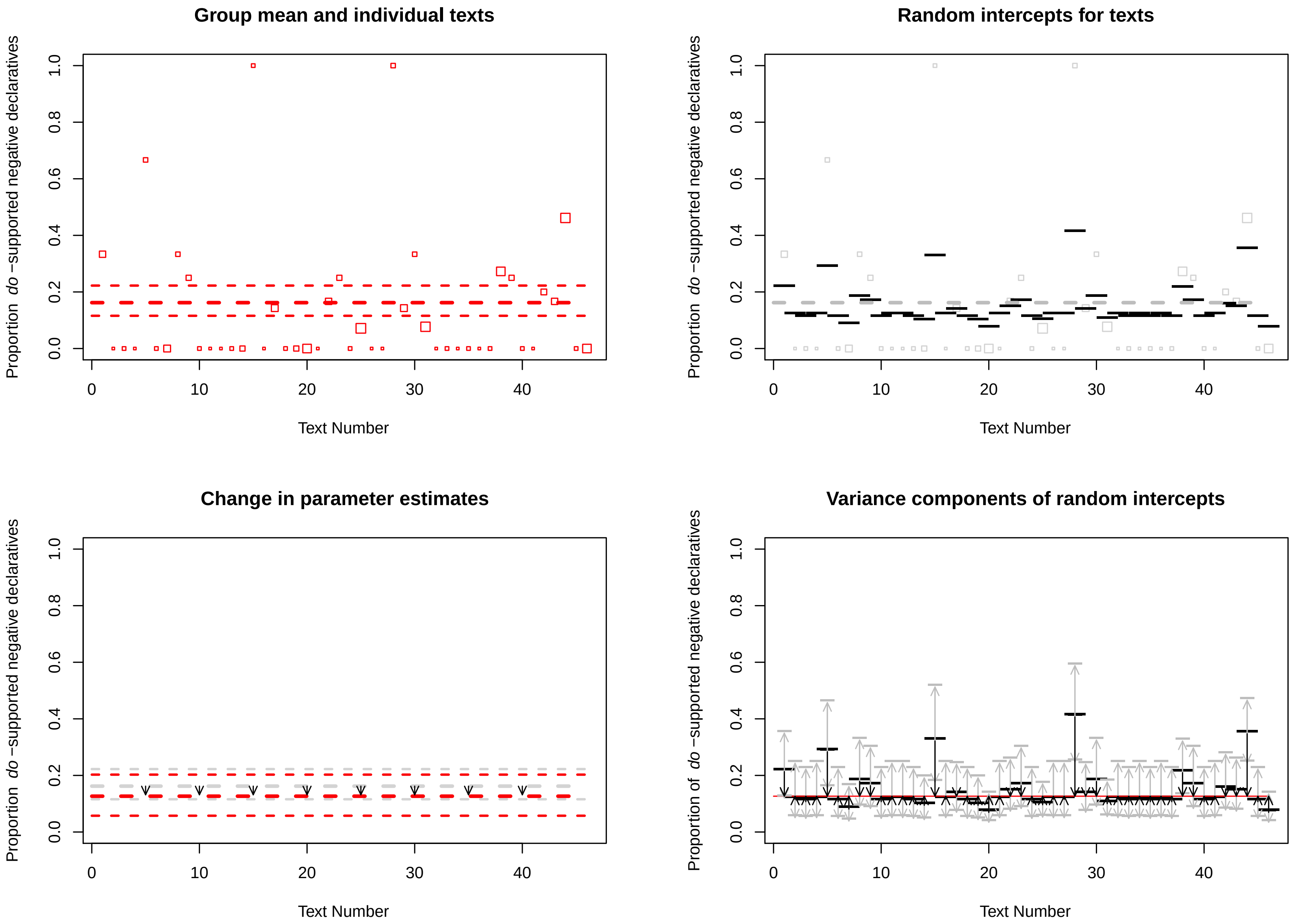



Figure 2.8: Illustration of the random text effect for the year 1900

(i) The top left panel of figure 2.8 represents a fixed-effects model fitted to the negation data for 1900 only. It returns an intercept of -1.64, which corresponds exactly to the proportion of innovative possessive *have*, i.e., 30 innovative cases out of a total of 185 examples = 16.2%. This year average is shown by the intercept of the thick, red, dashed line. The model also returns 0.1995 as the standard error of the intercept. That means that the true mean of innovative possessive *have* for the year 1900 can be expected to fall within the 95%-confidence interval -1.64 ± 1.96*0.1995, i.e., [11.6% – 22.2%]. The confidence interval is indicated by the thinner, red dashed lines around the mean. However, the data is actually composed of 46 individual texts. Those are represented in the graph as red boxes proportional in size to the number of examples they contain. There is considerable variation in the realization of possessive *have* between these texts. For example, the majority of texts, 28 or 61%, do not feature any innovative *do*-support structures at all; two texts are very innovative and have a proportion of 100% *do*-supported negative declaratives, etc. The single intercept in the model does not factor in the individual differences between these texts in any way. (ii) The top right panel of figure 2.8 illustrates a varying intercepts model for the negation data of the year 1900. It repeats for convenience the group mean at around 16% and the individual text boxes from the previous graph in light gray and additionally depicts random intercepts estimated for each of the 46 texts as thick, black lines. Texts that are considerably more innovative than the group as a whole are given a random intercept above the group mean, whereas texts that are more conservative fall below. Moreover, the intercepts of larger text boxes may have larger distances from the group mean; the intercepts of smaller text boxes remain closer to the group intercept. (iii) The bottom left graph of figure 2.8 shows the change in the parameter estimates resulting from the inclusion of the random 'Text' variable. Specifically, the point estimate for the group intercept changes from the earlier 16.2% of the fixed-effects model, indicated by the thick, gray line, to just 12.7% under the mixed-effects model, as shown by the thick, red line. The lower intercept more adequately reflects the fact that most of the 46 constituent texts are actually quite conservative. Thus, the population of American English texts from the year 1900 as a whole can be expected to cluster around a lower average of innovative possessive *have* than previously assumed. The standard error of the intercept is now estimated as 0.3417, which results in a 95%-confidence interval for the group average of [6.9% – 22.1%]. The range of the interval has widened considerably, from 10.6%, shown by the thinner gray lines around the fixed-effects group mean, to 15.2%, indicated by the thinner red lines around the mixed-effects group mean. The wider confidence interval indicates greater uncertainty in the point estimate of the group intercept and in effect makes it harder for the estimate of the parameter to reach statistical significance. (iv) Finally, the bottom right plot in figure 2.8 visualizes the calculation of the variance of the varying intercepts. The overall variance is made up of two variance components. On the one hand, the discrepancies between the random intercepts and the group mean vary among the different texts. This is shown by the dark arrows ranging between the black individual text intercepts and the red line of the mean for the year 1900. On the other hand, there is also variation among the standard errors of each of the varying intercepts themselves. The light arrows extending above and below the random intercepts illustrate the uncertainty in the estimates for each intercept. Both of these kinds of errors are squared, summed up and averaged over the total number of texts to yield the two variance components. Here, the variance of the discrepancies between the 46 random intercept point estimates and the mean is about 0.219. The variance of the uncertainty within each of the random intercept estimates themselves is about 0.655. The sum of the two variance components yields the variance of the random text intercepts returned by the mixed-effects model of about 0.874.

There are several problematic aspects of mixed-effects models. First, it is difficult to test for the significance of the random effects. In fact, I will refrain from testing the random text variable and merely evaluate the degree of variation within texts as a subjective intuition. Secondly, it is unclear how many degrees of freedom the random effect has. I will assume that the single variance parameter for the random text effect counts as 1 degree of freedom. Third, there are different ways of constructing null models, in particular, by different likelihood estimators and with or without the random effects structure. For simplicity's sake, I will simply use the null model from the previous fixed-effects models (even though this may not be entirely appropriate). Finally, the estimation of the parameters is substantially complicated in mixed-effects model, so that it can be hard to estimate their standard errors and to calculate confidence intervals. I will assume that the estimates returned by *R* are reliable (despite certain error messages).

Table 2.6 below presents the mixed-effects model that includes, in addition to the fixed effects discussed above, varying intercepts for every text.

```
formula = Have ~  Year + Genre + Year:Genre + (1 | Text),
       family = binomial, data = negation

Fixed effects:
                          Estimate   Std.Error   z-value   p
Intercept                 -110.6     1.498       -73.88    <0.001***
Year                      0.05729    0.0007726   74.15     <0.001***
Genre(non-fic→fic)        22.13      1.564       14.15     <0.001***
Year:Genre(non-fic→fic)   -0.01162   0.0008057   -14.42    <0.001***

Random effect:
Text, N=7,586
Variance of random intercepts:  1.411

Null deviance:         45792 on 33073 degrees of freedom
Residual deviance:     25701 on 33069 degrees of freedom
AIC: 25711
```

Table 2.6: Logistic Regression Model 3: mixed-effects model for negative declaratives

The fixed effects can be interpreted as in the previous model. The 'Year' coefficient predicts an increase in the log-odds of innovative *'do not have'* of 0.05729 for every one year advance in time. This is considerably faster than the time estimate without the random text effect. At this rate of change, the innovative variant of possessive *have* is modeled to rise from 1% to 99% of use in just 160 years, and in fact is predicted to have breached the 99% threshold in 2011. The 'Genre' coefficient indicates that the log-odds of *do*-supported negative declaratives are estimated to rise by 22.13 as the genre category is changed from non-fictional to fictional texts. This difference is comparable to the previous model fit. The interaction term between genre and time indicates that the estimated log-odds of innovative possessive *have* will increase by 0.01162 less for fictional than for non-fictional texts. Hence, the rate of change for fictional texts would only be 0.05729 -0.01162 = 0.04567. At this rate of change, it would take 201 years for the innovative form to spread from 1% to 99% of use, and would be predicted to reach the 99% level in 2038. Again, this difference in slope is assumed to follow from register fluctuations. The rate of change will appear slower in contexts that allow for the imitation of forms different from a speaker's actual grammar, for instance as a literary device, as a jocular expression etc. The three variables remain highly significant predictors for the realization of possessive *have* in the mixed-effects model.

The random 'Text' variable reports the number of individual texts aggregating the data, here 7,586, and the variance of the varying intercepts, here 1.411. This variance parameter seems to me to indicate substantial variation among texts. For example, the variance parameter predicts that texts will show with a certainty of 95% a proportion of innovative possessive *have* between [8.9% – 91.1%] during the midpoint of the change, where the overall probability of finding the innovative form is 50%. A text's percentage of innovative possessive *have* should fall within the 95%-confidence interval [1.1%– 53.4%] during an early stage of the change, where the average probability of the new form is just 10%, etc.. The wide ranges of the interval, 82.2% and 52.3% in the first and second scenario respectively, suggest considerable variability among texts.

The mixed-effects model is illustrated by the graph in figure 2.9. The small, gray boxes represent the proportion of innovative *have* in each individual text. The red and orange dots correspond to the same frequency binned into years for non-fictional and fictional texts respectively. The regression lines and their 95%-confidence bands plot the model predictions for the realization against time for each of the genre categories.

The model performs very well. It fits significantly better than a null model (where the null model is the same as for the preceding two models) (Likelihood Ratio Test, $\chi^2$=25087.14, df=4, $p<0.001$). The model fits well (Pseudo-$R^2_{marginal}$= 0.595 (before: 0.577), Pseudo-$R^2_{conditional}$= 0.717). The model's classification precision is very high (C-index=0.941 (before: 0.896), classification accuracy = 86.4% (before: 82.6%) vs. baseline: 52.1%).

The reasoning presented in this section will carry over to the following data evaluations as well. Specifically, the structure of the mixed-effects model for the negation data in table 2.6 will be repeated for the subsequent contexts. This will ensure comparability between the different data sets.

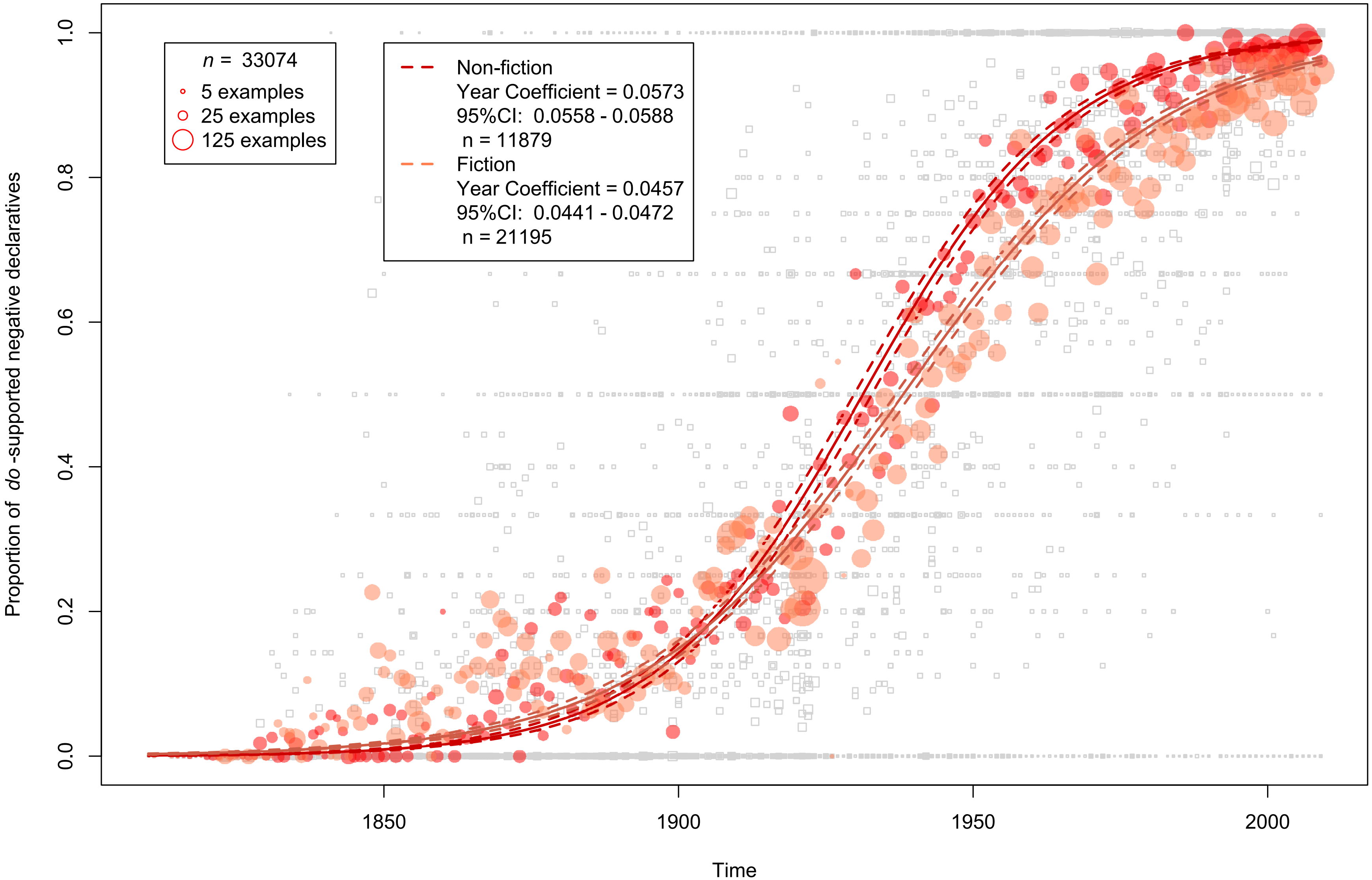



Figure 2.9: Illustration of the mixed-effects model for the negation context

## 2.3.4 Inversion

Next, I will turn to the inversion context. I collected and analyzed two separate data sets, for direct question and negative inversion, respectively, as outlined in the theoretical part under section 2.2.1.4.

### 2.3.4.1 Data Collection

The search queries for direct questions were put together in the following way. The dependent linguistic variable had two levels. The conservative variant showed a finite form of possessive *have* in front of a subject, *'have subject'*, while the innovative variant involved the dummy auxiliary *do* in front of the subject followed by possessive *have*, *'do subject have'*.

The search queries varied according to the following independent variables. First, I targeted subjects with a length of either 1 or 2 words. One-word subjects were defined as a set of part of speech labels, namely any kind of determiner, `[d*]`, any kind of indefinite pronoun, `[pn*]`, any kind of personal pronoun, `[pp*]`, as well as any kind of proper name `[np*]`. Two-word subjects were searched for by a combination of articles, determiners or possessive adjectives, `[at*]|[d*]|[app*]`, immediately followed by any kind of noun, `[n*]`.

Second, questions could occur with three kinds of polarity - they could be positive, involve high negation or low negation. Queries for positive questions did not contain a negator at all. Queries for low negation placed the word `not` immediately after the subject. Queries for high negation showed the contracted form of negation, `n't`, immediately before the subject. The position and form of the two negations were deduced from my grammar model, which cannot generate structures with a verb under `C` and full form *not* (see example (32)), but can do so effortlessly where negation is contracted (see examples (54) and (55)). Thus, I did not search for examples of non-contracted negation in pre-subject position and a brief investigation showed that the number of such structures in the COHA does in fact appear to be vanishingly small.

Third, adverbs, labeled `[r*]`, were allowed to occur in the structure. An adverb could occur immediately after the subject (to find, e.g., *Did she never have a chance?*) or, where present, after low negation (to find, e.g., *Did she not ever have a chance?*). If an adverb was found, its lexeme was manually added as an additional variable to the data set as well.

Fourth, a finite form of conservative *have* or innovative *do* was placed immediately before the targeted subject, or, where present, before high negation. As before, it could be annotated for three different inflectional forms, 'Present' for the searches involving the finite forms `have` and `do`, 'Present 3$^{rd}$ singular' for `has` and `does` and 'Past' for `had` and `did`.

Fifth, the likelihood of actually finding a possessive *have* structure was increased by also looking for an element that probably indicated the beginning of a possessed element in the last position of the search string. In general, those were the same object types as for the negation data set (see table 2.2). However, searches for bare singular and bare plural nouns were omitted. The reason was that those queries would have found a great many ambiguous results for the conservative variant with one-word subjects, for which it would have been plausible that the supposed subject really formed a constituent with the element targeted as the bare noun. The examples in (141) illustrate ambiguous strings of that kind.

(141) a. have [$_{\text{targeted one-word subject}}$ most] [$_{\text{targeted bare singular}}$ reason ]
(vs. have [$_{\text{DP}}$ most reason ] )

b. has [$_{\text{targeted one-word subject}}$ this] [$_{\text{targeted bare singular}}$ power ]
(vs. has [$_{\text{DP}}$ this power ] )

Considering such structures would have imposed an unreasonable strain on the manual correction process. Hence, there remained only 8 levels of the 'Object type' variable: `the`, `a`, `an`, strong quantifiers, weak quantifiers, numerals, demonstratives and possessives.

Moreover, these 8 object types were replaced by a completely different set of *wh*-object elements in search queries that attempted to find examples in which the object itself was a fronted *wh*-element. For those cases, the *wh*-object elements were put in first position whereas the last position was occupied by a question mark, `?`, or an adverbial expression followed by a question mark, `[r*]|[md] ?`. The *wh*-object element was searched for by the strings `what|which` targeting the items *what* and *which* in their function as *wh*-pronouns as explained in the discussion around example (51), or by the strings `what|which *` to find the same items functioning as *wh*-determiners with a one-word complement as illustrated in example (52). Furthermore, I looked for the strings `how many|much` to find object questions with those words in first position, as well as for `how many|much *` to retrieve examples of object questions starting with *how many* or *how much* followed by a one-word complement. The resulting *wh*-object elements are presented in table 2.7 below.

<table>
<tr><th>Variant name</th><th>Targeted object</th><th>Search term</th></tr>
<tr><td>WhArg1a</td><td>wh-pronoun</td><td><code>what|which ... ?</code></td></tr>
<tr><td>WhArg2a</td><td>wh-determiner</td><td><code>what|which * ... ?</code></td></tr>
<tr><td>How1a</td><td>how + 1 word</td><td><code>how many|much ... ?</code></td></tr>
<tr><td>How2a</td><td>how + 2 words</td><td><code>how many|much * ... ?</code></td></tr>
<tr><td>WhArg1b</td><td>wh-pronoun, final adverb</td><td><code>what|which ... [r*]|[md*] ?</code></td></tr>
<tr><td>WhArg2b</td><td>wh-determine, final adverb</td><td><code>what|which * ... [r*]|[md*] ?</code></td></tr>
<tr><td>How1b</td><td>how + 1 word, final adverb</td><td><code>how many|much ... [r*]|[md*] ?</code></td></tr>
<tr><td>How2b</td><td>how + 2 words, final adverb</td><td><code>how many|much * ... [r*]|[md*] ?</code></td></tr>
</table>

Table 2.7: The levels of the 'object type' variable in object questions

Sixth, I targeted four different types of questions. Yes-no questions were found by queries that placed in initial position a set of elements that are likely to identify the beginning of a sentential unit. This set included punctuation signs, namely full stop, question and exclamation mark, colon and semicolon, comma, and inverted commas, as well as the conjunctions *and*, *but*, *or*, and proper names (intended as vocatives), `.|?|;|!|:|"|,|and|but|or|[np*]`. These queries placed an object element in last position. Adjunct questions were retrieved with queries that showed in first position one of the simplex *wh*-adjuncts, *why*, *where*, *when*, *how* or *whence*. Again, the last slot of these queries was filled by an object element. Next, argument questions showed in initial position the special variants of the object type variable shown in table 2.7. Finally, I also searched for questions whose initial *wh*-element likely had a long-distance dependency with a more deeply embedded function. Queries for this question type began with one of the *wh*-elements *whom*, *who*, *what* or *which* and also had an object element in last position. These queries would pick up, among others, examples of *wh*-items extracted from non-finite complements of the possessed element, as in (142), or initial *wh*-elements that functioned as complements of prepositions, either pied-piped or stranded, (143).

(142) a. and [whom] have I the pleasure of addressing _ ?
WhiteMule 1937

b. [Whom] do I have the pleasure of addressing _ please ?
Whirlwind 1986

(143) a. and [on whom] have I any influence _ ?
ImpressionsEnglish 1947

b. [Which] do you have most affection for _ ?
JohnWardPreacher 1888

The data set is thus based on a total of 2 (dependent variable) * 2 (subject length) * 3 (polarity) * 2 (adverb) * 3 (inflection) * 8 (object type) * 4 (question type) = 2,304 search queries. 64 of those search queries could not be run because they exceeded the length limit on search queries imposed by the COHA interface. Those were replaced by an additional 64 search queries, which looked for object questions involving *wh*-phrases with a length of 3 or 4 words, as in (144), as well as for yes-no questions with 3, 4 and 5-word subjects, as in (145).[7]

(144) a. [What general bodily sensation] have you?
AllenHouse 1860

b. [What other old objects] do you have?
NewMoon 1994

(145) a. and has [our system of prescientific psychology] the right to open the door to such glittering epigrams
PsychologySocial 1914

b. Does [anyone in the social sciences] have a clue to the nature of collecting?
Home-PsychTheSocial 1983

[7]The exact search queries are saved on the Dissertation DVD under `Chapter 2 - Possessive Have/04 Inversion Context/Questions/Question Queries.xlsx`.

Some examples of search queries are the following. The search query `why|when|where|how|whence has [at*]|[d*]|[app*] [n*] every|each|all|most|least|both|neither` targeted the conservative variant of possessive *have* in an adjunct question with a two-word subject, positive polarity, no adverb, 3rd person singular inflection and a strongly quantified object; the search query `.|?|;|!|:|"|,|and|but|or|[np*] did n't [d*]|[pn1*]|[pp*]|[np*] [r*] have that|those|this|these` targeted the innovative variant of possessive *have* in a yes-no question, with a one-word subject, high contracted negation, an adverb, past inflection and a demonstrative object, etc.

After all the data had been collected, I attempted to correct all precision errors. The most common type of error in my question data resulted from a failure to separate between the targeted subject and object functions so that perfect *have* could intrude. For example, (146a) illustrates a targeted object element that turned out to be a part of the subject (*we each*, similarly *all the*, *such a* etc.). The search query that retrieved example (146b) involved a definite description as the targeted object, but it really found a relative clause dependent on the subject. The underlinings highlight the strings that were falsely matched by a search query.

(146) a. Why haven't we each paid the $200?
ChristCentury 2007

b. How had the expressions the prisoners used come into being?
FieldBrokenStones 1950

Similarly, example (147) presents a case of a targeted one-word subject, *many*, that actually formed a constituent with the object. The example was found because it included a topic-dropped subject.

(147) . Had many a wonderful time at the Cabin.
Harpers 1936

I also deleted verb-first conditionals dependent on *had*, (148).

(148) . Had she a dagger she would have cut my heart out.
TamedByYourDesire 2002

There were also a number of unwanted instances of small clause constructions typically involving causative *have*.

(149) a. Why don't you have some of your own mob do this job?
Empire 1951

b. how has she the churl subdued!
MatildaDiShabran 1834

A sentence's question type was changed manually if the correction process revealed that it more adequately instantiated a different category. For example, the sentence below was picked up as a yes-no question. However, it contains the *wh*-item *why* in front of the relevant search space. Its question type was therefore re-annotated as an adjunct question.

(150) *Question type change from* yes-no *to* wh-adjunct *question*:
Why, oh, why, do they never have any pieces?"
UnderCountrySky 1916

Finally, some fronted *wh*-element turned out to function as subjects. Subject questions are not included in my data set because they never display *do*-support. Such sentences were therefore deleted.

(151) a. which things have most beauty?
RoderickHudson 1876

b. Which ports have the greatest lagging?
ManualPhysical 1909

The final data set comprises 9,319 instances of direct possessive *have* questions.[8] The examples below illustrate the relevant variation with some near minimal pairs.

[8]The question data set is saved on the Dissertation DVD under `Chapter 2 - Possessive Have/04 Inversion Context/Questions/Question Data.xlsx`.

(152) a. *conservative variant:* '*have* subject'
**Have you a** Bible in this office?
EighthDay 1967

b. *innovative variant:* '*do* subject *have*'
**Do you have a** doctorate in linguistics?
ColorPurple 1982

(153) a. *conservative variant:* '*have* subject'
**Why haven't you a** bell?
Play:Greatness 1921

b. *innovative variant:* '*do* subject *have*'
**Why don't you have a** car?
WheelLove 1970

The data set for negative inversion structures was assembled in a different way. Instead of designing search queries around a large number of independent variables, I only varied the search term for the initial negative constituent and manually filtered the results.

The dependent variable could take on the following two forms. For the conservative variant, the initial negative constituent was immediately followed by a form of *have*, 'neg *have*.' For the innovative variant, the initial negative constituent was immediately followed by a form of *do,* and the collocate search function was used to find the word *have* up to 9 words to the right of this string, 'neg *do ... have*.' I did not find any examples of the conservative variant with a distance between *have* and the first word of the possessed element greater than 8 words. Hence my principle of symmetric searches was obeyed - the search queries gave the conservative and innovative variants an equal chance of being retrieved.

The only independent variable for this context was the kind of negative initial constituent. A total of 11 different elements were searched for. Specifically, the initial string could be (1) a punctuation sign followed by the word *never*, (2) a punctuation sign followed by the word *never* followed by another adverb, (3) a preposition followed by *no* followed by a noun, (4) a preposition followed by *no* followed by an adjective, noun or numeral, followed by a noun, (5) the string *not only*, (6) a punctuation sign, followed by *not* followed by any adverb other than *only*, (7) *nor*, (8) *neither*, (9) *neither* followed by a noun, (10) *no* followed by an adverb, and (11) *nowhere*. The following three examples illustrate. I underlined the strings that were matched by a specific search query.

(154) *never* (negative initial constituent type 1)

a. <u>. Never has</u> a lover more grace than when he deplores the pangs he suffers
LastDuelInSpain 1832

b. <u>. Never did</u> a man <u>have</u> a more devoted adherent in his wooing than did I
HeartHappyHollow 1904

(155) complex negative PP (negative initial constituent type 4)

a. He felt that <u>at no previous moment</u> had he so much at stake as now;
DukesPrizeStory 1848

b. <u>On no other continent does</u> the cycle <u>have</u> such an extreme impact.
NaturalHist 1993

(156) *nowhere* (negative initial constituent type 11)

a. <u>Nowhere had</u> that despotic leader more violent and unscrupulous partisans than there.
Century 1886

b. <u>NOWHERE did</u> the Industrial Revolution <u>have</u> greater | effect than in the United States of America
AnOutlineHistory 1968

Hence, I conducted a total of 2 (dependent variable) * 11 (negative initial constituent) = 22 search queries. The inflectional forms of *have* and *do* were manually added as a variable to the data set.[9]

As a trade-off for the small number of search queries, I had to spend a lot of time on the correction of precision errors. Search queries for the conservative variant frequently returned two kinds of unwanted results. Firstly, the sentences involved not possessive but perfect auxiliary *have*. Hence, the search string was followed by a subject and a past participle rather than a subject and an object, as in (157) below.

[9] The negative inversion search queries are stored on the Dissertation DVD under `Chapter 2 - Possessive Have/04 Inversion Context/Negative Inversion/Negative Inversion Queries.xlsx`.

(157) Not only have these views **been** largely received by the Church of England, but ...
NewEngYaleRev 1866

Second, the element after the finite verb frequently functioned not as a subject but as an object, as in (158).

(158) ... when neither side had a new proposal to offer on an exchange of Korean war prisoners.
NYT-Reg 1952

After the elimination of such and various other, less common false hits, the data set was left with 1,040 examples.[10] The following near minimal pair serves as an illustration of the relevant variation in the negative inversion data.

(159) a. *conservative variant:* 'neg *have*'
**nor have I any intention** of overthrowing the Government of the United States by violence!
Harpers 1942

b. *innovative variant:* 'neg *do ... have*'
**nor do I have any intention** of communicating with him.
NYT-Reg 1948

The combined data set for the inversion context thus consists of a total of 9,319 (questions) + 1,040 (negative inversion) = 10,359 examples. In the next section, I will analyze the diachronic development of possessive *have* in this data.

#### 2.3.4.2 Results

I evaluated the inversion data with a mixed-effects model structured analogously to the last model of the previous context. That is to say, the model includes the 'Year' variable, the difference between fiction and non-fiction and the interaction between those two factors as fixed effects, as well as varying intercepts for every text as a random effect. The result is shown in table 2.8 below.

```
formula = Have ~  Year + Genre + Year:Genre + (1 | Text),
       family = binomial, data = inversion
```

Fixed effects:

| | Estimate | Std.Error | z-value | $p$ |
|---|---|---|---|---|
| Intercept | -116.0 | 2.107 | -55.07 | $<0.001^{***}$ |
| Year | 0.05976 | 0.001085 | 55.09 | $<0.001^{***}$ |
| Genre(non-fic→fic) | 27.65 | 2.092 | 13.22 | $<0.001^{***}$ |
| Year:Genre(non-fic→fic) | -0.01422 | 0.001078 | -13.19 | $<0.001^{***}$ |

```
Random effect:
Text, N=4,255
Variance of random intercepts:  1.304

Null deviance:         14170 on 10358 degrees of freedom
Residual deviance:     8248  on 10354 degrees of freedom
AIC: 8258
```

Table 2.8: Logistic Regression Model 4: mixed-effects model for the inversion context

The graph in figure 2.10 illustrates this model. It should be interpreted in exactly the same way as the graph of the parallel model for the negation context in the previous section. The data points and regression lines for non-fictional and fictional texts are represented in blue and purple color respectively.

[10] The negative inversion data set is saved on the Dissertation DVD under `Chapter 2 - Possessive Have/04 Inversion Context/Negative Inversion/Negative Inversion Data.xlsx`.

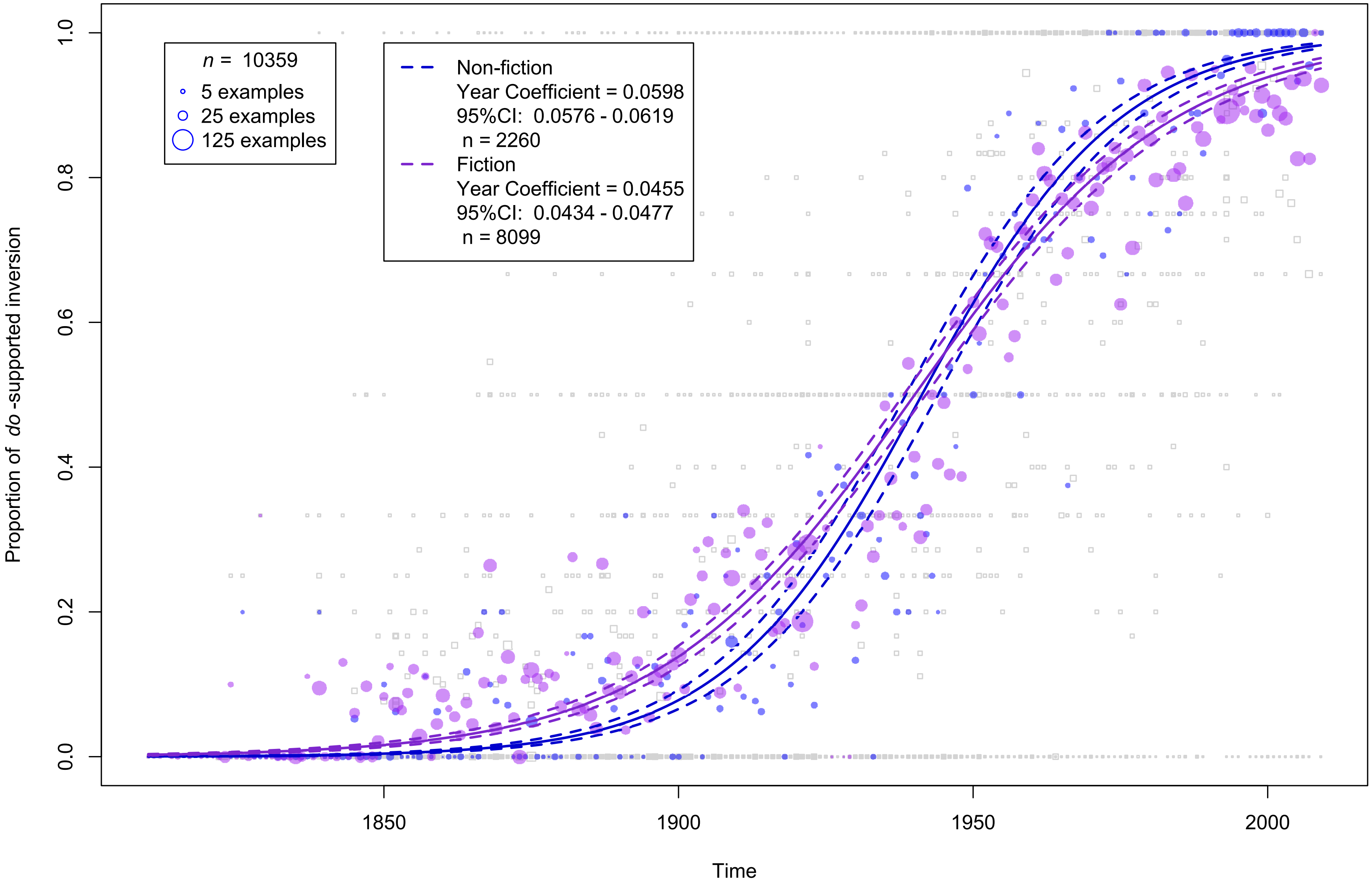



Figure 2.10: Illustration of the mixed-effects model for the inversion context

The 'Genre' coefficient estimates an increase in the log-odds of *do*-supported inversion by 27.65 as the genre category is changed from non-fiction to fiction. The 'Year' coefficient in this model predicts for every one year advance in time an increase in the log-odds of innovative possessive *have* of 0.05976 in non-fictional, and of 0.05976 - 0.01422 = 0.04554 in fictional texts. At these rates of change, the innovative form would spread from 1% to 99% of use within a time span of 153 years for non-fictional and 201 years for fictional texts, breaking the 99% threshold of use in 2018 and 2041 respectively. This finding shows that, as hypothesized, *do*-supported possessive *have* inversion structures do indeed become prevalent during the late Modern American English period. All fixed effects of this model are highly significant predictors for the realization of possessive *have* in the inversion context. The inversion data comes from a total of 4,255 texts, whose random intercepts show a variance of 1.304. Following the same reasoning as for the negation context, this parameter may be taken to indicate substantial variability in the proportion of innovative *have* among individual texts.

The mixed-effects model for the inversion context performs in a satisfactory way. It fits significantly better than a null model (where the null model is assumed to be an intercept-only model) (Likelihood Ratio Test, $\chi^2$=7848.31, df=4, $p$<0.001). The overall model fit is good (Pseudo-$R^2_{marginal}$= 0.597, Pseudo-$R^2_{conditional}$= 0.711). The model's classification precision is very high (C-index=0.944, classification accuracy = 87.0% vs. baseline: 56.8%).

The significant difference in slope between non-fiction and fiction can once again be explained as an effect of intrusive forms from vernacular or archaic registers into the non-authorial speech of fictional writings for literary purposes. The excerpts in (160) illustrate portrayals of vernacular, lower-class speech in otherwise high-register fictional texts from the early stages of the change.

(160) a. "**Did you ever have the rheumatis?**" inquired Mistress Debora, pathetically, and rubbing her shoulder as she spoke. [...] "It's pesky bad, I can tell you [...] I've had it now these four weeks e'en-a-most, I've tried hards'-lard, that Dame Jenkins recommended to me, - she that lives by the sign of the stump-tailed-bull, [...] Debby, says she, and says I what? Where's Thankful? says she, and says I, Thankful's down stairs"
NixsMateAnHistorical 1839

b. I is loss fum de cumpny, but dis is de ferry, and I spec dey'll soon come. **But didn't we have a good time** las' night in Buffalo? Dem dar Buffalo gals make my heart flutter, dat dey did. But, tanks be to de Lord, I is got religion. I got it las' night in de meetin.' Before I got religion, I was a great sinner; [...]. But now I is a conwerted man; I is bound for hebben;
Escape 1861

c. "Bliaff was a brate bid man, an' Dave was brate little man, an' Bliaff said,' Come over here'n an' I'll eat you up,' an' Dave said,' I ain't fyaid of you.' [...] an' Dave took Bliaff's sword an' sworded Bliaff's head off, an' made it all bluggy, an' Bliaff runned away. [...] I'd like to hear' bout Ferus." "Who?" "Ferus; don't you know?" "Never heard of him, Budge." "Why – y – y –!" exclaimed Budge; "**didn't you have no papa** when you was a little boy?"
HelensBabies 1880

Example (160a) is a part of a long monologue by a character described as a "garrulous old woman." Her non-standard features include first person singular agreement in *-s* (*says I*), and colloquial expressions (*pesky bad, e'en-a-most* < *even almost*, meaning 'nearly'). Example (160b) is uttered by a slave named 'Cato.' Orthography indicates distinct features of African-American Vernacular English for his passage. For example, he seems to pronounce the first sound of words that have word-initial inter-dental fricatives in the standard variety as alveolars (e.g., *this, the, thanks* as *dis, de, tanks*), reduces final consonant clusters (e.g., *and*, *last*, *expect* as *an'*, *las'*, *spec*), and realizes standard intervocalic *v* as *b* (*heaven* as *hebben*) (AAVE phonological features number 5, 1 and 12 in Rickford 2001: 4-5). Example (160c) comes from a child called Budge. His non-standard forms are manifest, among others, in infantile confusion between certain consonants (e.g., *bid* for *big*, *bluggy* for *bloody*) or regularizations of irregular past tenses (*runned* for *ran*). The fact that these excerpts also include *do*-supported possessive *have* inversion structures, shown in bold-face, can thus plausibly be attributed to the same portrayal of vernacular speech.
Conversely, the excerpts in (161) are likely examples of character representations in historical settings or of old speakers' diction in fictional writings from the late stage of the change.

(161) a. "Read the papers then." "I hate them. The old ink comes off on me hands. **Haven't you any little thing of ye own** to show me? **Have ye any sketches**?" He shifted around in the bed, as if he would give anything to be out of it.
SouthernRev 1992

b. Begin with her face. It is thin, you say, but well formed? **Has she not the snub nose** and round cheeks of so many Sunderland girls whose raw ancestors tramped down from Scotland or washed ashore lo those many centuries ago from pork-fed Saxony?
DressLodger 2001

c. "I can not feel a heartbeat." Phaedra's voice was surprisingly calm given the terror rising in her throat. "**Have you a doctor close at hand**, Mr. Kemble?" "Yes, yes, just round the corner."
TemptedAllNight 2009

Example (161a) is spoken by a very old man in a nursing home. Example (161b) is taken from the historical novel *The Dress Lodger* set in Sunderland, England during the cholera epidemic of 1831. Finally, example (161c) occurs in the same historical romance as the sentence in (140b), which was used to illustrate the fictional portrayal of old-fashioned parlance in the negation context. The bold-faced instances of conservative possessive *have* in these excerpts may thus be viewed as a part of the archaizing literary style of these texts.

It is quite likely that there are other, substantial genre differences in the inversion data. For example, low negation in questions and negative inversion in general sound quite formal and stilted and might thus more commonly appear in non-fictional writings. Questions with high, contracted negation, on the other hand, probably occur predominantly in colloquial fiction. Moreover, positive yes-no questions are far more common than *wh*-questions or questions involving negation overall. Disparate distributions of this kind might have interesting consequences for the measurement of the diachronic development of possessive *have* as well. However, since such factors cannot be compared across the different syntactic contexts, I will not explore them here in any more detail. Instead, I shall now turn to the third syntactic context to be investigated, the distribution of adverbs and floating quantifiers.

### 2.3.5 Adverbs and Floating Quantifiers

In this section, I will present the data collection for possessive *have* sentences that co-occur with several of the adverbs identified as relevant for my purposes in section 2.2.1.5 or with subject-dependent floating quantifiers as explained at the same place. Subsequently, I will analyze the data statistically taking into account the fact that my formal grammar model allows for high adjunction of these adverbs.

#### 2.3.5.1 Data Collection

The adverb data was collected in the following way. Search queries varied with respect to the relative order between an adverb and the verb *have*. This yielded two variants of the dependent variable. The conservative variant placed the adverb after the verb, '*have* adverb.' The innovative variant showed the adverb in pre-verbal position, 'adverb *have*.'

Next, the searches coded for three independent variables. As before, every example was annotated for the inflectional form of the finite verb *have*, 'Present' for `have`, 'Present 3rd singular' for `has` and 'Past' for `had`. Further, I once again required the presence of one of ten possible object elements following *have* to increase the chance of finding instances of possessive *have*. These object elements were the same as in the preceding contexts, i.e., *the*, *a*, *an*, a strong quantifier, a weak quantifier, a numeral, a demonstrative, a singular noun, a plural noun or a possessive adjective. Third, the search queries were applied to 25 different sentence modifiers that met the requirements laid out in the theoretical section on adverbs in 2.2.1.5. This set comprised the following items: some modal adverbs (with evidential or epistemic flavor), namely (1) *necessarily*, (2) *certainly*, (3) *surely*, (4) *probably*, (5) *possibly*, (6) *really* and (7) *actually*, the frequency adverbs (8) *always*, (9) *ever*, (10) *often*, (11) *frequently*, (12) *usually*, (13) *sometimes*, (14) *rarely*, (15) *seldom* and (16) *never*, the time point adverb (17) *now*, (18) *soon* and (19) *then*, the aspectual adverbs (20) *no longer* (terminative), (21) *still* (continuative), (22) *already* (anterior) and (23) *suddenly* (inceptive), as well as (24) concessive *nevertheless* and (25) additive *also*.

Thus, I carried out a total of 2 (dependent variable) * 3 (inflection) * 10 (object type) * 25 (adverb lexeme) = 1,500 search queries to collect the adverb data.[11] For example, the search query `have often that|this|those|these` targeted examples that involve the conservative variant, present tense inflection, a demonstrative object, and the adverb *often*. The query `also had [app*]` retrieved sentences that were supposed to instantiate the innovative variant with past tense inflection, a possessive adjective nominal and the adverb *also*.

After all tokens had been accumulated, I manually corrected precision errors. I deleted sentences that involved pre-modification of the adverb. Such complex adverb phrases were not retrieved for the conservative

[11] The search queries are listed for each of the 25 investigated adverbs in separate excel files under the tab "QUERIES." These files are stored on the Dissertation DVD under `Chapter 2 - Possessive Have/05 Adverb Context/Adverbs Separately`.

variant and would thus have biased the data set towards innovative forms thereby violating the principle of symmetric searches. Examples of offending structures of this kind are shown below.

(162) a. The reader **most probably has** some slight remembrance of Mienheer Jacobus Kip's pointed though brief remarks
BuccaneersARomance 1827

b. I also **very soon had** many opportunities to know that the goodness of his heart was equal to the acuteness of his judgment
PartnersInScience 1969

Further, some forms of *have* turned out to instantiate unwanted perfect auxiliaries. The targeted object element would typically function as an adjunct in such cases. Below and in the following examples, the underlined elements represent the string falsely matched by a particular search query.

(163) it has still more facilitated inter-migration.
Arena Volume4 1849

I discarded all cases of non-finite *have*, such as the following.

(164) a. Finally, small presses, literary magazines, university reviews, do still have a place in my writing life.
Harpers 1988

b. so Molly's girl might always have a dainty frame for her unusual beauty.
DrumsJeopardy 1920

c. If you want to really have a scramble, just let hog prices drop three or four dollars.
SatEvePost 1958

d. Congress has never had the opportunity of proving itself
NewRepublic 1958

Finally, I removed sentences in which the targeted object seemed to function as a subject of a subsequent predicate in a small clause construction. Such instances would usually involve a form of causative *have*.

(165) a. The upshot was that I soon had her telling me of a play she had recently seen.
RiseDavidLevinsky 1917

b. Baxter also had her students excavate a couple of present-day backyards
HistoryToday 1970

The final data set has a size of 28,271 examples.[12] As before, I illustrate the relevant variation with some near minimal pairs.

(166) a. *conservative variant:* '*have* adverb'
He **had always** a strong sense of his responsibility and duty to his Creator.
Home 1835

b. *innovative variant:* 'adverb *have*'
I **always had** a strong sense of him being lucky.
NYT 1988

(167) a. *conservative variant:* '*have* adverb'
I **have certainly** no desire that this discussion should be unreasonably protracted.
ReportDebatesProceedings 1864

b. *innovative variant:* 'adverb *have*'
Now I **certainly have** no desire to defend Kant's theological views
NewRepublic 1986

[12]The complete data set can be found on the Dissertation DVD under `Chapter 2 - Possessive Have/05 Adverb Context/Adverbs Combined/All Adverbs Data.xlsx`. The data is also listed in the independent excel files for each of the 25 adverbs under the tab "DATA", see `Chapter 2 - Possessive Have/05 Adverb Context/Adverbs Separately`.

(168) a. *conservative variant:* ‘*have* adverb’
France **has no longer** any regard for the rights of Prussia or the Confederation
NYT-Reg 1860

b. *innovative variant:* ‘adverb *have*’
Great Britain **no longer has** the strongest fleet
Atlantic 1921

The data set for the floating quantifiers, *all*, *both* and *each*, was assembled in a very similar manner. The main difference was that the dependent variable relied on floating quantifiers and not on adverbs as diagnostics of the verb position, yielding the two levels ‘*have* floating quantifier’ for the conservative and ‘floating quantifier *have*’ for the innovative variant. Likewise, the manual correction process was in essence carried out in the same way as for the adverbs. In fact, the search queries and the correction process for *each* were identical to the ones for adverbs. The queries for *all* and *both*, however, had to be restricted more carefully to avoid large numbers of false hits in the search output, and the manual correction process included additional steps to limit the hits to unambiguous examples of quantifier float only. I will now explain the modifications for these two quantifiers.

Firstly, sentences with the floating quantifiers *all* and *both* were not sampled by search queries for finite verbs agreeing in 3rd person singular, `has` and `does`. The reason was simply that these floating quantifiers necessarily modify subjects involving plural or at least collective nominal forms.

Next, the word *all* easily forms a constituent with nominal phrases that are bare or introduced by elements such as the definite article, determiners, numerals or possessive adjectives. For example, the following sentences are ambiguous - the word *all* could be interpreted as a subject-dependent floating quantifier, or as a quantifier of the object. The two interpretations are clarified by unambiguous paraphrases.

(169) a. *all* + *the*
My father’s guests have **all the** news
Victory 1955
i. ‘All of my father’s guests have the news’ (floating quantifier)
ii. ‘My father’s guests have the entire set of news’ (object quantification)

b. *all* + numeral
These recipes have **all three** of these characteristics.
ChildDigest 1999
i. ‘All of these recipes have three characteristics’ (floating quantifier)
ii. ‘These recipes have the three familiar, exhaustive characteristics’ (object quantification)

c. *all* + bare singular
Methodist bishops have **all power** at all times.
RevReviews 1903
i. ‘All Methodist bishops have power at all times.’ (floating quantifier)
ii. ‘Methodist bishops have every power at all times’ (object quantification)

To reduce the risk of contamination from this kind of ambiguity, I changed some of the object elements co-occurring with possessive *have*. I required the words *same*, *one* or *common* to follow the definite article and the demonstrative. The hits produced by these queries were then counted as instances of quantifier float. An example is shown below.

(170) sister, son, and the son of the half-brother by the same father, had **all the same** malformation.
PhysicalLifeWoman 1896

The search queries involving numerals were pruned to the terms `one|1` only. Again, the sentences these queries retrieved were accepted as examples of floating quantifiers.

(171) Mankind have **all one** common origin;
HillsShatemuc 1852

I could not find practicable modifications for the search queries involving bare singular nouns, bare plural nouns or possessive adjectives and thus did not conduct searches for these problematic object elements at all.

Third, the word *both* imposes similarly rigid constraints on its interpretability as a floating quantifier. Specifically, many instances of *both* followed by an unquantified plural or coordinated object are ambiguous between quantifier float and object quantification. This is illustrated in (172). In fact, the irrelevant, second interpretation seems to be favored quite strongly.

(172) ... that the Communists **had both** 57-and 75-mm. recoilless rifles.
LookingGlassWar 1965
i. 'Both of the Communists had 57-and 75-mm recoilless rifles.' (floating quantifier)
ii. 'The Communists had two things, namely 57-and 75-mm recoilless rifles.' (object quantification)

I tackled the problem of potential distortion of the actual distribution of floating quantifiers by such structures in two ways. On the one hand, I did not conduct searches for bare singular and bare plural objects at all because the identification of the small number of relevant hits from these object elements would have imposed an incommensurate burden on the manual correction process. On the other hand, I manually removed all sentences with *both* and an unquantified plural and / or coordinated object.

Finally, I eliminated all sentences with *all* or *both* that also featured the personal pronoun subjects *we, you* or *they* during the manual correction process. The rationale for this decision was the following. I believe that pronouns, but not common nouns, are likely to form a constituent with the quantifiers *all* and *both*. Consequently, a large number of the strings 'pronominal subject - *all*/*both* - *have*' might not be parsed with adjunction of a floating quantifier to `I'` by rule (111a) or to `VP` by rule (111b) but rather with a complex subject containing a pronominal head and a subsequent quantificational modifier. The inclusion of such structures would then have biased the dataset towards the innovative variant. There are several arguments in support of this decision.
(i) The oblique forms *us* and *them* can clearly form constituents with *all* and *both* in object position but full `DP`s cannot.

(173) a. She saw {[them all] / [us all]}.
b. * She saw [my friends all].

It is not implausible to assume that a similar constraint operates on subjects.
(ii) The relative frequency of the structure 'pronoun + quantifier', such as *we all*, is far greater than the relative frequency of 'plural noun + quantifier', such as *my friends all*, in the COHA. Specifically, 25.6% of all instances of a pronominal subject and the quantifier *all* up to 5 words to its right show the subject and the quantifier immediately adjacent to each other, but only 8.1% of all instances of plural nouns do. Similarly, the same difference for the quantifier *both* is striking, with a proportion of 41.9% for pronouns and 7.6% for plural nouns respectively. The raw data for these findings is presented in table 2.9.

| Structure | Quantifier up to 5 words to the right | Of those, immediately adjacent |
|---|---|---|
| *we, you, they ... all* | 150,852 | 38,577 (25.6%) |
| plural noun ... *all* * | 175,906 | 14,163 (8.1%) |
| *we, you, they ... both* | 25,875 | 10,823 (41.9%) |
| plural noun ... *both* * | 35,085 | 2,674 (7.6%) |

* consideration of first 1,000 plural nouns returned

Table 2.9: Frequency of [pronoun + Q] vs. [full noun + Q] structures in the COHA

The discrepancy can be explained by assuming that *all*/*both* form a constituent with personal pronouns more commonly than with plural nouns.
(iii) The second person plural pronoun *y'all*, typical of Southern American English, appears to have originated from the string *you* + *all*, perhaps under the influence of the Scots-Irish phrase *ye aw* (Montgomery 1992). The pronoun and the quantifier are thus likely to have formed a constituent before the grammaticalization process. It is not unreasonable to assume that other pronoun and quantifier clusters have the same underlying structure.
(iv) There are also some constituent tests that may suggest that nominative personal pronouns but not plural nouns can form constituents with *all* and *both*. However, standard constituency tests are very difficult to apply to pronominal subjects because of case requirements (e.g., *it*-clefts disprefer nominative foci), the availability of alternative parses and the relative inflexibility of subjects in general. Nevertheless, the following two tests may be illuminating. First, non-restrictive relative clauses may modify a pronoun + quantifier head, but not a plural noun + quantifier head. This follows if the element modified by the relative clause must form a constituent. In the following example, I rate the modification of the pronoun + quantifier constituent with a question mark because my native speaker informants found that such structures sound somewhat dated.

(174) a. ? And [we all], who are followers of this religion, will live forever after we die.
b. * And [my friends all], who are followers of this religion, will live forever after they die.

Second, a subject can marginally be isolated in verb phrase ellipsis structures with subject-auxiliary inversion. My native speaker informants did in fact report a contrast in the acceptability of personal pronoun + quantifier vis-á-vis plural noun + quantifier subjects isolated in this manner. Again, this can be explained if the former but not the latter string can easily form a constituent. In the example below, I use a question mark to show that some of my informants rated the pronoun + quantifier constituent as clunky or literary.

(175) a. ? In case of a war, soldiers will suffer, as will [we all].
b. * In case of a war, soldiers will suffer, as will [my friends all].

All in all, then, there seems to be sufficient reason to believe that personal pronouns + *all*/*both* can form complex subjects and to therefore eliminate those structures from my dataset.

The floating quantifier data is based on a total of 2 (dependent variable) * 3 (inflection) * 10 (object type) for the quantifier *each*, plus 2 (dependent variable) * 2 (inflection) * 7 (object type) for the quantifier *all*, plus 2 (dependent variable) * 2 (inflection) * 8 (object type) for the quantifier *both* = 120 search queries.[13] For instance, the search `has each an` targeted sentences with the conservative variant involving the floating quantifier *each*, 3$^{rd}$ person singular present tense inflection and an object introduced by the indefinite article *an*, whereas the query `all had one|1` was meant to retrieve examples of the innovative variant involving the floating quantifier *all*, past tense inflection, and an object beginning with the numeral *one*.

After the manual correction had been completed, there remained 594 examples of possessive *have* structures with floating quantifiers in my data set.[14] The following near minimal pairs illustrate the variation in the relative order between *have* and floating quantifiers.

(176) a. *conservative variant:* '*have* floating quantifier'
Fichte, Schelling, Hegel, **had each** his own system, though they have been called transcendentalists.
NewEngYaleRev 1843
b. *innovative variant:* 'floating quantifier *have*'
His mother, Fraulein Schlote and Miss Letitia Lamb **each had** her own accent and intonation
LastPuritan 1936

(177) a. *conservative variant:* '*have* floating quantifier'
Our arts, stages, steam-boats, hotels, and shops, **have all** a share of his property
NewEngMag 1831
b. *innovative variant:* 'floating quantifier *have*'
the educational institutions, the health-care system, the social-services system, and the government **all have** a part in the success of corporations
USCatholic 1996

(178) a. *conservative variant:* '*have* floating quantifier'
the red limit and the violet limit **have both** the same luminous intensity
TreatiseOnForces 1845
b. *innovative variant:* 'floating quantifier *have*'
Arcturus and Capella **both have** the same magnitude
ExploringDistant 1956

I will now investigate the diachronic development of possessive *have* structures with adverbs and floating quantifiers.

#### 2.3.5.2 Results

I shall first present a preliminary analysis of the development of the relative order between adverbs and possessive *have* based on a simple logistic regression model that predicts the two possible syntactic structures from the time variable only. The structure of the regression equation is thus analogous to the first logistic regression model I discussed for the negation context in table 2.3. The result is shown in table 2.10 below. The following plot in figure 2.11 is a graphical representation of that model.

[13] The search queries for floating quantifiers are saved on the Dissertation DVD under `Chapter 2 - Possessive Have/05 Adverb Context/Floating Quantifiers/Floating Quantifiers Queries.xlsx`.
[14] The floating quantifier data set can be found on the Dissertation DVD under `Chapter 2 - Possessive Have/05 Adverb Context/Floating Quantifiers/Floating Quantifiers Data.xlsx`.

```
formula = Have ~  Year, family = binomial, data = adverb

              Estimate     Std.Error     z-value     p
 Intercept    -50.77       0.7997        -63.50      <0.001***
 Year         0.02740      0.0004216     64.99       <0.001***

 Null deviance:         24625 on 28270 degrees of freedom
 Residual deviance:     18705 on 28269 degrees of freedom
 AIC: 18709
```

Table 2.10: Logistic Regression Model 5: possessive *have* against time in adverb context

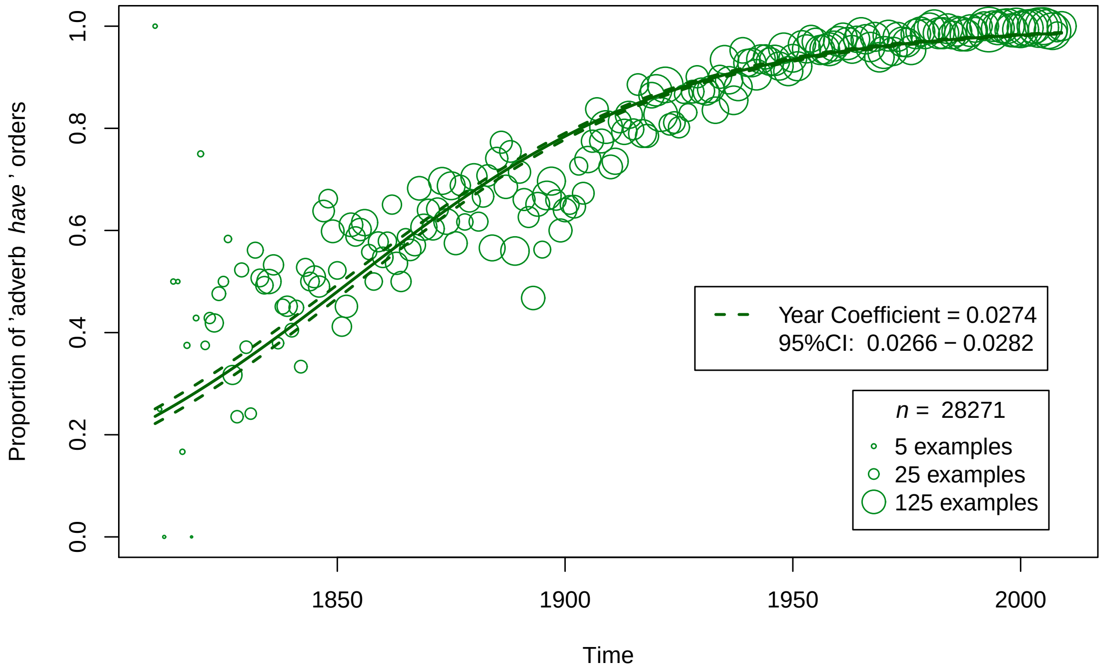



Figure 2.11: Development of possessive *have* in the adverb context

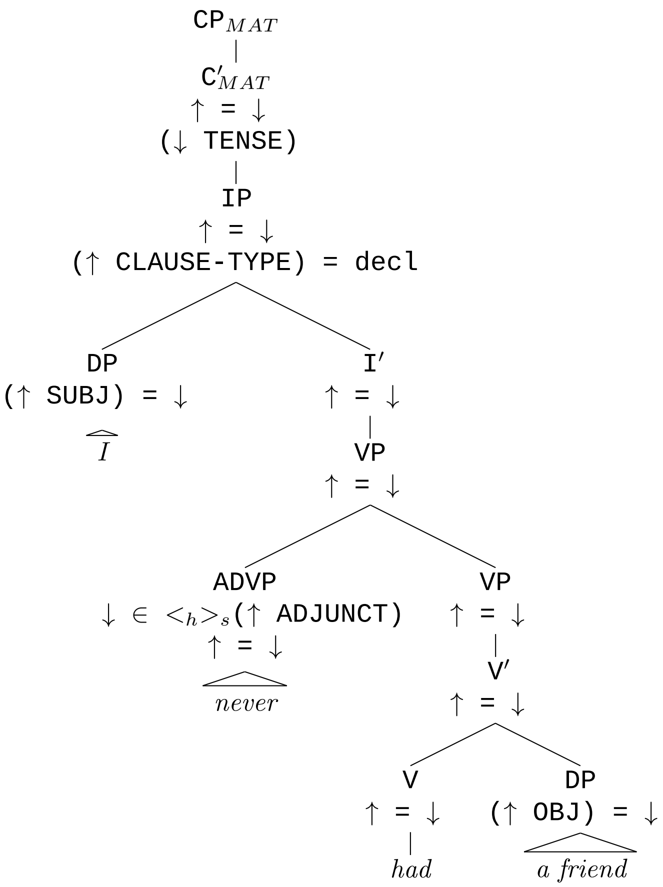
The statistical analysis reveals an unmistakable increase in the order 'adverb *have*' during the late Modern American English period. The data thus confirms the predicted change in that direction.

However, the analysis also suggests that the rate of change is markedly different between the adverb context and the preceding negation or inversion contexts. The slope of the regression curve is much flatter for the former than for the latter. For example, the Year-coefficient in the above model is estimated at 0.0274, a rate at which the innovative word order would be expected to rise from 1% to 99% of use in 335 years. In contrast, the Year-coefficient in the structurally equivalent model for the negation context in table 2.3 was estimated at a considerably higher value of 0.0414, and this rate would cause the innovative word order to increase to the same degree within 223 years. In other words, the change appears to proceed substantially more slowly when measured with adverbs than with negation or inversion structures.

This observation does not constitute a refutation of the Constant Rate Hypothesis, but is entirely expected under my model assumptions, as I shall show now. By necessity, my study cannot use adverbs that unfailingly occur after auxiliaries, but rather relies on items that typically occur after auxiliaries but can also marginally be positioned in front of them (see section 2.2.1.5, e.g., example (77)). My model parses such structures as low adjunction to VP for the dominant and as high adjunction to I′ for the minority pattern respectively (see the rules in (90)). Consequently, the surface order 'adverb *have*' must be regarded as structurally ambiguous. It may be parsed with the innovative variant of possessive *have* and an adverb in its preferred VP-adjoined position. This is the parse that the adverb diagnostic relies on because the adverb is assumed to function as a boundary marker between higher, conservative and lower, innovative *have*. At the same time, the structure can also be parsed with conservative possessive *have* provided that the adverb is placed in its less common pre-auxiliary position. This second parse is uninformative for the distinction between the higher and the lower verb position. Example (179) illustrates the structural ambiguity for the earliest instance of a pre-verbal adverb in my data set, *I never had a friend*.

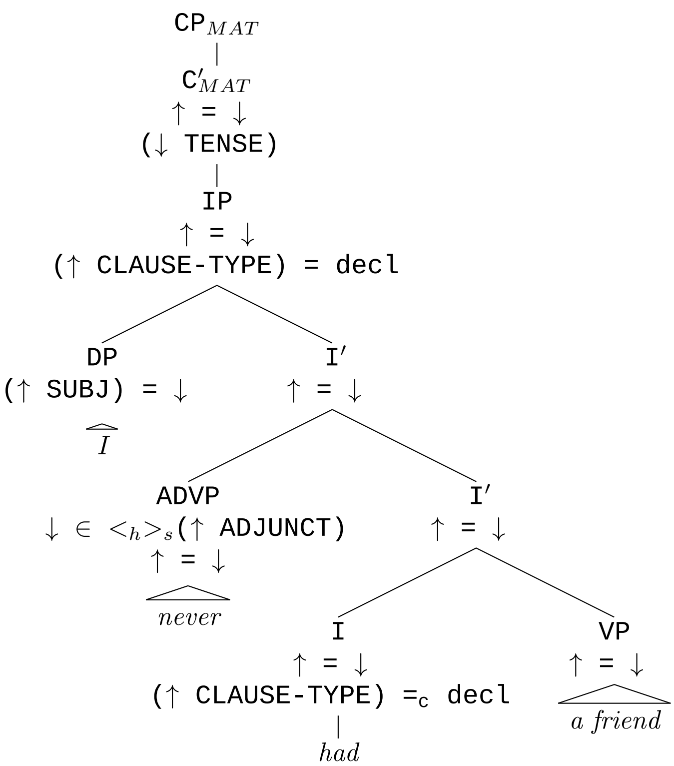
(179) I never had a friend
GordianKnot 1810

a. *diagnostic parse - low-adjoined adverb identifies innovative* have *to its right*




$$\left[\begin{array}{ll}
\text{PRED} & \text{‘HAVE}\langle(\text{SUBJ}),(\text{OBJ}_\theta)\rangle\text{’} \\
\text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{‘PRO’}\\ \text{PERS} & 1 \\ \text{NUM} & \text{sg}\\ \text{PRON-TYPE} & \text{personal}\end{array}\right] \\
\text{OBJ}_\theta & \left[\begin{array}{ll}\text{PRED} & \text{‘FRIEND’}\\ \text{PERS} & 3 \\ \text{NUM} & \text{sg}\\ \text{NOUN-TYPE} & \text{common-count}\\ \text{DEF} & \text{-}\end{array}\right] \\
\text{ADJUNCT} & \left\{\left[\begin{array}{ll}\text{PRED} & \text{‘NEVER’}\end{array}\right]\right\} \\
\text{TENSE} & \text{past} \\
\text{VFORM} & \text{fin} \\
\text{CLAUSE-TYPE} & \text{decl}
\end{array}\right]$$

b. *uninformative parse - high-adjoined adverb is compatible with conservative* have *to its right*




$$\left[\begin{array}{ll}
\text{PRED} & \text{‘HAVE}\langle(\text{SUBJ}),(\text{OBJ}_\theta)\rangle\text{’} \\
\text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{‘PRO’}\\ \text{PERS} & 1 \\ \text{NUM} & \text{sg}\\ \text{PRON-TYPE} & \text{personal}\end{array}\right] \\
\text{OBJ}_\theta & \left[\begin{array}{ll}\text{PRED} & \text{‘FRIEND’}\\ \text{PERS} & 3 \\ \text{NUM} & \text{sg}\\ \text{NOUN-TYPE} & \text{common-count}\\ \text{DEF} & \text{-}\end{array}\right] \\
\text{ADJUNCT} & \left\{\left[\begin{array}{ll}\text{PRED} & \text{‘NEVER’}\end{array}\right]\right\} \\
\text{TENSE} & \text{past} \\
\text{VFORM} & \text{fin} \\
\text{CLAUSE-TYPE} & \text{decl}
\end{array}\right]$$

Examples of pre-verbal adverbs will thus confound the adverb test to some degree. To see this, it helps to imagine that 100% of all adverbs were adjoined to `I'`. In this scenario all examples would be categorized as innovative pre-verbal adverbs, the data points for every year would lie at 1, and there would be no measurable change. If, on the other hand, 100% of all adverbs were adjoined to `VP`, only innovative instances of possessive *have* would be identified as such, the data points for every year would follow the trajectory of the actual frequency of *have* in the lower verb position, and the adverb test would constitute a perfect diagnostic to measure the change. As it happens, the true rate of adverbial high adjunction lies in between these two extremes. Hence, some but not all innovative 'adverb *have*' sentences should really be regarded as conservative. Those instances will distort the model estimates and pull up the regression line, thus seemingly slowing down the change.

The question then becomes if it is in fact true that adverbs can adjoin high to `I'` in conjunction with possessive *have* structures so that high adverbial adjunction could obfuscate the true rate of change. There are in fact a great many constructions that demonstrate quite unequivocally that an adverb can occur in front of conservative possessive *have*. (i) Direct negation may unambiguously identify conservative possessive *have*. A co-occurring pre-verbal adverb thus appears to be quite clearly adjoined high. Here and in the subsequent examples, my model parse is sketched with labeled bracketing diagrams.

(180) *post-verbal negation forces conservative* have *+ high adverb*

a. We **certainly have not** the solidarity and endurance of the Jews
... [$_{\mathrm{I'}}$ certainly [$_{\mathrm{I'}}$ have [$_{\mathrm{NEGP}}$ not ... ]]]
Atlantic 1927

b. They **probably haven't** a cutting-torch at all.
... [$_{\mathrm{I'}}$ probably [$_{\mathrm{I'}}$ haven't ... ]]
Aliens 1965

(ii) Possessive *have* can be identified as conservative by a following adverb or floating quantifier. An additional, preceding adverb is thus very likely to be high-adjoined.

(181) *post-verbal adverb or floating quantifier forces conservative* have *+ high adverb*

a. The poor Indians **now have often** reason to rejoice
... [$_{\mathrm{I'}}$ now [$_{\mathrm{I'}}$ have [$_{\mathrm{VP}}$ often [$_{\mathrm{VP}}$... ]]]]
NaturalistInNicaragua 1847

b. Sir Henry comes forward to indicate a policy which will have the united support of the party, and **apparently has already** the approval of the country.
... [$_{\mathrm{I'}}$ apparently [$_{\mathrm{I'}}$ has [$_{\mathrm{VP}}$ already [$_{\mathrm{VP}}$... ]]]]
Nation 1905

c. they **still had each** a strong following
... [$_{\mathrm{I'}}$ still [$_{\mathrm{I'}}$ had [$_{\mathrm{VP}}$ each [$_{\mathrm{VP}}$... ]]]]
Atlantic 1893

(iii) The conservative variant of possessive *have* can also be determined by a verb-phrase ellipsis structure licensed after a finite form of *have* that reconstructs a meaning based on a possessive *have* source clause. If an additional adverb occurs before finite *have* in this environment, it can only be parsed with high adverbial adjunction in my model.

(182) *direct verb phrase ellipsis forces conservative* have *+ high adverb*

a. You have a flair, Townsend. You **really have** $\Delta$.
... [$_{\mathrm{I'}}$ really [$_{\mathrm{I'}}$ have ]]
Play:Veronique 1962

b. but if anybody had a reason , he **certainly had** $\Delta$
... [$_{\mathrm{I'}}$ certainly [$_{\mathrm{I'}}$ had ]]
Play:StreetScene 1928

(iv) Finally, the presence of conservative possessive *have* is not necessitated but at least suggested by the presence of coordinated or otherwise parallel phrases with an adverb in pre-auxiliary position. An adverb in front of possessive *have* is thus once again likely to be adjoined high.

(183) *parallel adverb+auxiliary cluster suggests conservative* have *+ high adverb:*

a. But, he **never was** respected; much less, revered; and, **never had** any permanent influence;
... $[_{I'}$ never $[_{I'}$ was ... ]] ... and $[_{I'}$ never $[_{I'}$ had $[_{VP}$... ]]]
RandolphANovel 1823

b. a woman **never can** obtain rank by merit, therefore **never has** reason to be proud of it.
... $[_{I'}$ never $[_{I'}$ can ... ]] ... therefore $[_{I'}$ never $[_{I'}$ has $[_{VP}$... ]]]
LoversVows 1814

It now remains to be shown that sentences with high adverb placement are indeed a major confounding factor causing the reduced effect of time on the variation in possessive *have* in the adverb context by comparison with the negation or inversion contexts. The following prediction can be tested to discover if my explanation can in fact be assumed to be correct: If the rate of change in the adverb context is lowered by a substantial number of 'adverb *have*' sentences involving conservative possessive *have* and a high-adjoined adverb, then there should be a correlation between (A) the rate of change measured by individual adverbs and (B) an independent measure of their propensity to adjoin high. I tested this hypothesis as follows.

(A) The rate of change measured by an individual adverb was simply operationalized as the Year-coefficient returned by a logistic regression run separately on the data for each of the 25 adverbs of this study. The structure of the regression equation was kept constant for all adverbs; it included, again, the dependent word order variable and time as the only predictor. Unfortunately, the sample size of each adverb fluctuates widely in my data set. It ranges from a very low number of 26 for the uncommon adverb *possibly* to a very large number of 4,913 for the high frequency item *still*. Since the accuracy of the parameter calculations depends on the sample size, the point estimates of the Year-coefficient were found to be very imprecise for low-frequency adverbs. For example, the rate of change for the adverb *possibly* was estimated to fall within the 95%-confidence interval [-0.0053 – 0.0374] yielding a wide range of 0.043, whereas the rate of change for the adverb *still* was predicted to lie with a 95% probability somewhere between [0.0349 – 0.0403] showing a much more narrow range of just 0.005. I therefore decided to consider only those items in my data set with a sample size larger than 500 observations. This left the following set of 14 adverbs for consideration: *then, usually, often, probably, certainly, really, ever, no longer, already, always, now, never, also* and *still.* Despite the drastic reduction from 25 to just 14 adverb types, more than 90% of the tokens could still be retained in the data set (26,460 out of 28,271 examples).

(B) An adverb's propensity to adjoin high was measured through a High Adjunction index. For this index, I counted all instances of a personal pronoun, `[pp*]`, or noun, `[nn*]`, as a target for the subject followed by a particular adverb before and after the modal auxiliaries *may, might, can, could, shall, should, will, would* and *must* as well as before and after finite forms of the primary auxiliary *be*, *is, are, was* and *were* in the COHA. Assuming that these word forms instantiate heads of category `I`, the proportion of adverbs in pre-auxiliary position should give a rough measurement of an adverb's overall tendency to occur high.[15] Table 2.11 exemplifies the determination of the High Adjunction index for the adverb *never*, which was found to be 16.1%.

| auxiliary form | `[pp*]\|[nn*]never` *aux* | `[pp*]\|[nn*]` *aux* `never` | % before auxiliary |
|---|---|---|---|
| *may* | 42 | 1,414 | 2.9 |
| *might* | 9 | 1,044 | 0.9 |
| *can* | 1,480 | 5,617 | 20.9 |
| *could* | 2,687 | 6,809 | 28.3 |
| *shall* | 614 | 4,155 | 12.9 |
| *should* | 580 | 3,387 | 14.6 |
| *will* | 1,322 | 8,151 | 14.0 |
| *would* | 2,008 | 11,643 | 14.7 |
| *must* | 12 | 945 | 1.3 |
| *is* | 139 | 2,923 | 4.5 |
| *are* | 94 | 1,674 | 5.3 |
| *was* | 1,824 | 7,921 | 18.7 |
| *were* | 324 | 2,459 | 11.6 |
| TOTAL | 11,135 | 58,142 | 16.1 |

Table 2.11: Example of the High Adjunction index for the adverb *never*

[15] The High Adjunction index for each of the 25 investigated adverbs can be found under the tab "INDICES" of the individual excel files stored on the Dissertation DVD under `Chapter 2 - Possessive Have/05 Adverb Context/Adverbs Separately`.

Interestingly, the value of *never*'s High Adjunction index, 16.1%, is remarkably close to previous estimates for this adverb based on different methodologies (e.g., Kroch 1989, Ecay 2015: 57). This may suggest that the probability of a particular adverb to occur in pre- or post-auxiliary position is comparatively stable across time and genres and hence generalizes to a wide range of different English corpora. More research is needed to produce evidence for or against this speculative point.

In a final step, I correlated each of the 14 adverbs' rates of change with their High Adjunction indices. The result is summarized in figure 2.12. The graph shows the log-odds of the proportion of high adjunction on the y-axis and the increase of the log-odds of finding innovative possessive *have* per year on the x-axis. The size of the data points for each adverb is proportional to the number of tokens they contribute. Gray error bars around the data points represent the uncertainty in the estimate of their rates of change in the form of 95%-confidence intervals. A red dashed line indicates the strength of association between rate of change and high adjunction, which is measured as the Pearson's product moment correlation coefficient, $r$.

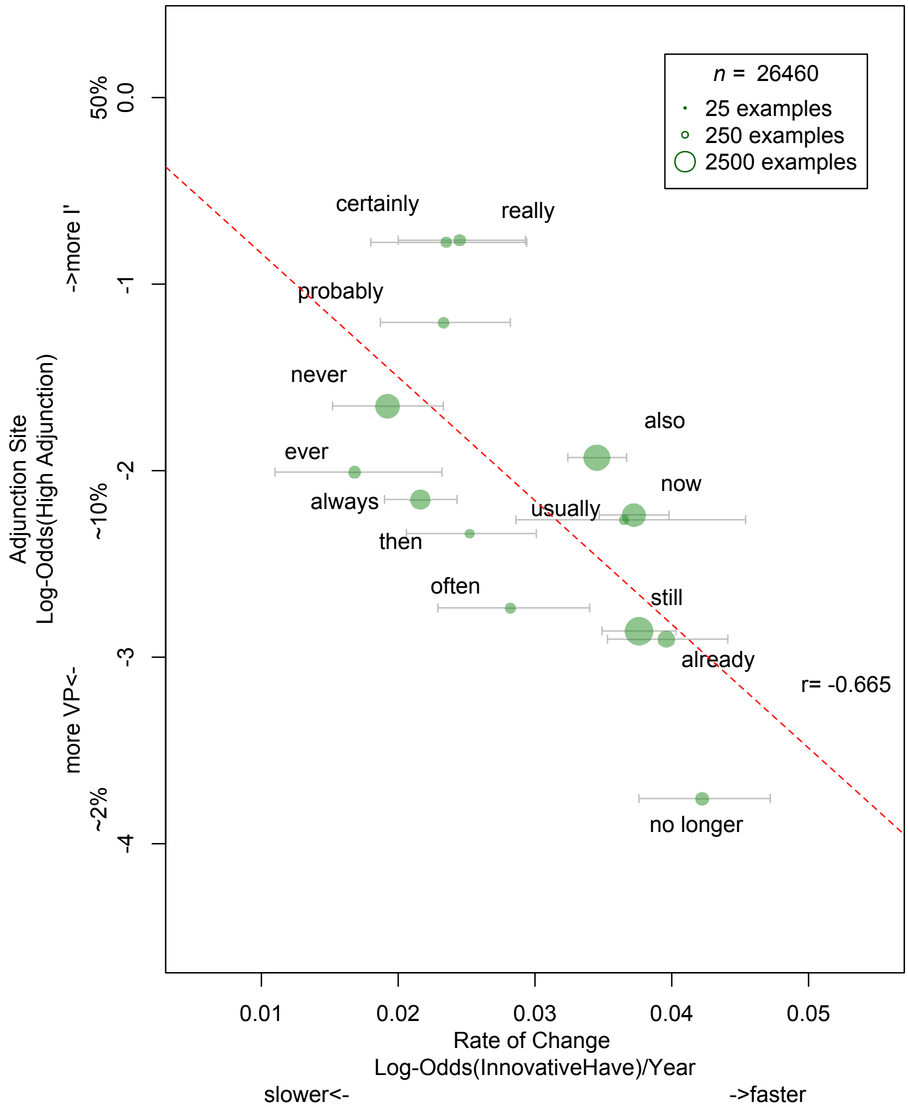



Figure 2.12: Correlation between an adverb's rate of change and its propensity for high adjunction

There appears to be a relatively strong linear relationship between the rate of change in possessive *have* as measured with a particular adverb and the adverb's tendency to occur before auxiliaries ($r$=-0.67, df=12, $p$<0.01). The more commonly an adverb is adjoined high, the slower the change it measures appears to be. The association between the rate of change and high adjunction can be explained with the mechanisms outlined above but is likely to remain mysterious otherwise: Conservative possessive *have* is modeled as distributionally indistinguishable from Present-Day English auxiliaries by virtue of sharing their syntactic category, `I`, so that those adverbs that are commonly placed before Present-Day English auxiliaries also commonly occur before conservative possessive *have*. Adverbs that are more likely to adjoin high will more frequently result in a parse that is uninformative for the status of possessive *have*. Therefore, adverbs that are more frequently positioned before auxiliaries will seemingly slow down the change in possessive *have* to a greater degree. Such being the case, this finding lends support to the hypothesis that high adverbial adjunction is a major reason for the relatively low rate of change in possessive *have* observed in the adverb context.

In order to measure the change in possessive *have* as accurately as possible in the adverb context, one can consider only those adverbs that are least likely to adjoin high and thus have the highest probability of indicating the VP-boundary. (An alternative would be to somehow adjust the estimated slopes based on a High Adjunction index, similar to the procedure in Kroch (1989) or Ecay (2015)). I decided to retain adverbs whose High Adjunction indices showed a value of about 5% or less. This criterion singled out only three adverbs from figure 2.12, *no longer*, *already* and *still*. I added the floating quantifier data to the three adverbs under the assumption that they, too, are predominantly adjoined to VP. I will collectively refer to these diagnostics as "low adjuncts" or "(high probability) VP-adjuncts."[16]. Table 2.12 presents some core characteristics of the low adjunct data set. It shows the sample size of each diagnostic element. In total, there are now 8,314 examples. This is less than the number of tokens in the previous negation and inversion contexts and much less than the 28,271 adverb sentences originally collected, but still constitutes a substantial data set that should allow measuring the rate of change of possessive *have* reasonably accurately. I also indicate the High Adjunction indices of the three adverbs. Finally, the table shows the rate of change for each of the diagnostics when independently regressed against time as the only independent variable as well as the 95%-confidence interval of the estimate. The four elements seem to behave relatively similarly as shown by similar values of the Year-coefficient of 0.0422, 0.0396, 0.0376 and 0.0399 for *no longer*, *already*, *still* and floating quantifiers respectively. (Ideally, one might include random slopes and intercepts for each lexical item in the model of the adverb data. I decided to omit such complications because it would not easily be possible to compare these random variables across all contexts.)

| diagnostic | *n* | High Adjunction index | rate of change | 95%-confidence interval |
|---|---|---|---|---|
| *no longer* | 1,080 | 2.3% | 0.0422 | [0.0376 – 0.0472] |
| *already* | 1,727 | 5.2% | 0.0396 | [0.0353 – 0.0441] |
| *still* | 4,913 | 5.4% | 0.0376 | [0.0349 – 0.0403] |
| floating quantifiers | 594 | - | 0.0399 | [0.0336 – 0.0469] |
| TOTAL | 8,314 | | | |

Table 2.12: Core characteristics of the low adjuncts data set

The adverb phrases *no longer*, *already* and *still* as well as floating quantifiers may be regarded as the Modern English high frequency items with the greatest diagnostic value for the onset of verb phrases. I would like to recommend the use of these items for future studies investigating the change of verb positions. One would have to ensure that these elements are not right-adjoined to the verb phrase as they would not be informative about the boundary between I and VP in this case. In effect, that means that the adjunct would have to be followed by additional verb phrase material, such as an object of a transitive verb, or the possessed element of possessive *have* in the present study. Even then it would still remain possible for an adverb to be parsed as adjoined to the right edge of VP provided that the verb phrase material was postposed, e.g., by heavy NP-shift. For example, the following sentence was included as a conservative instance of possessive *have* in my data set, but could, in principle, be analyzed as a structure with ambiguous *have*, a right-adjoined adverb, and postposition of the possessed element, as informally sketched by an arrow. The availability of a parse along these lines may explain the fact that some of my native speaker informants judge this sentence to be very marginally acceptable in Present-Day English.

(184) *sketch of unlikely but available postposition parse for post-verbal adverbs*
Lord Edward [$_{\text{VP}}$ [$_{\text{VP}}$ **has** _ ] **already**] [a friend]

Orphans 1818

Fortunately, postposition of this kind is likely to occur so infrequently as to introduce merely a small amount of noise into my data, and could be controlled for to some degree by restricting the length of the object in future studies. I therefore believe that *no longer*, *already*, *still* and floating quantifiers are superior to other

[16]The low adjunct data set can be found on the Dissertation DVD under `Chapter 2 - Possessive Have/05 Adverb Context/Adverbs Combined/Low Adjuncts Data.xlsx`.

VP-diagnostics. For instance, the adverb *never* has often been used to investigate changing verb positions (i.e., "the loss of verb movement" in derivational terms) in the history of English (e.g., Ellegård 1953, Ecay 2015, Haeberli and Ihsane 2016), but suffers particularly dramatically from contamination caused by high adjunction. This element may therefore be better suited as a rough estimator for the upper bound of the slope than as a precise measurement for the rate of change in verb placement. On the other hand, *never* has the advantage of being immune to ambiguities caused by right-adjunction or postposition.

I will now fit a mixed-effects logistic regression model to the new VP-adjuncts data. Its structure is identical to the final model of the negation data in table 2.6 and to the model of the inversion data in table 2.8, i.e., it includes time, genre and their interaction as fixed effects and varying intercepts for every text as a random variable. The result is shown in table 2.13 below.

```
formula = Have ~  Year + Genre + Year:Genre + (1 | Text),
       family = binomial, data = low_adjunct

Fixed effects:
                          Estimate   Std.Error   z-value   p
Intercept                 -99.6      4.163       -23.93    <0.001***
Year                      0.05287    0.002184    24.20     <0.001***
Genre(non-fic→fic)        10.60      5.657       1.87      0.0611
Year:Genre(non-fic→fic)   -0.00532   0.002973    -1.79     0.0738


Random effect:
Text, N=4,087
Variance of random intercepts:  1.911

Null deviance:        7128 on 8313 degrees of freedom
Residual deviance:    4144 on 8309 degrees of freedom
AIC: 4154
```

Table 2.13: Logistic Regression Model 6: mixed-effects model for high probability VP-adjuncts

The results of the logistic regression can be summarized as follows. The model returns an estimate of the Year-coefficient of 0.0529. As the genre category is changed from non-fiction to fiction, the value decreases slightly to 0.0529 - 0.0053 = 0.0476. As expected, the slopes are now substantially greater than before, which corroborates the view that the diagnostics selected here do indeed have a high probability of indicating the VP-boundary. Specifically, the above rates of change would predict a spread of 'low adjunct *have*' orders from 1% to 99% of use within 174 and 193 years for non-fictional and fictional texts respectively. The 'Year' variable is statistically significant in this model. This confirms the predicted development from post-verbal to pre-verbal placement of VP-adjuncts relative to possessive *have*. Neither the difference in the intercept nor the difference in the slope between the two genre categories constitute a significant predictor in this model. Nevertheless, I decided to report the model structure as it is to ensure comparability with the analogous models for the other contexts. The innovative form of possessive *have* seems to become generalized considerably earlier when measured with adverbs than when measured with negation or inversion structures. Specifically, the adverb model predicts crossing the 99% threshold of use of the innovative form in 1969 for fictional and 1972 for non-fictional writings. The completion of the change in the adverb contexts thus appears to chronologically precede the predicted future endpoint of the change in the negation or inversion contexts by about 40-70 years. The low adjunct data is taken from a total of 4,087 texts, whose random intercepts have a variance of 1.911. This parameter indicates even greater variability in the proportion of innovative possessive *have* among texts than in the negation or inversion data set.

The model quality seems to be reasonably high. It fits significantly better than a null model (where the null model is an intercept only model) (Likelihood Ratio Test, $\chi^2$=4297.11, df=4, $p$<0.001). The model fit to the data is good (Pseudo-$R^2_{marginal}$= 0.524, Pseudo-$R^2_{conditional}$= 0.700). The model's classification precision is high (C-index=0.967, classification accuracy = 92.5% vs. baseline: 84.7%).

The graph in figure 2.13 illustrates the model. It follows the same conventions for the representation of the combined year averages, individual texts and rates of change for non-fictional and fictional texts as before. There are two particularly noteworthy points. First, the regression curve shows a long tail towards the end of the change. This indicates the early completion of the change in the adverb context. Second, the light green and dark green regression lines for fiction and non-fiction run very parallel and proximal to each other. This reflects the statistically insignificant difference between the two genre categories.

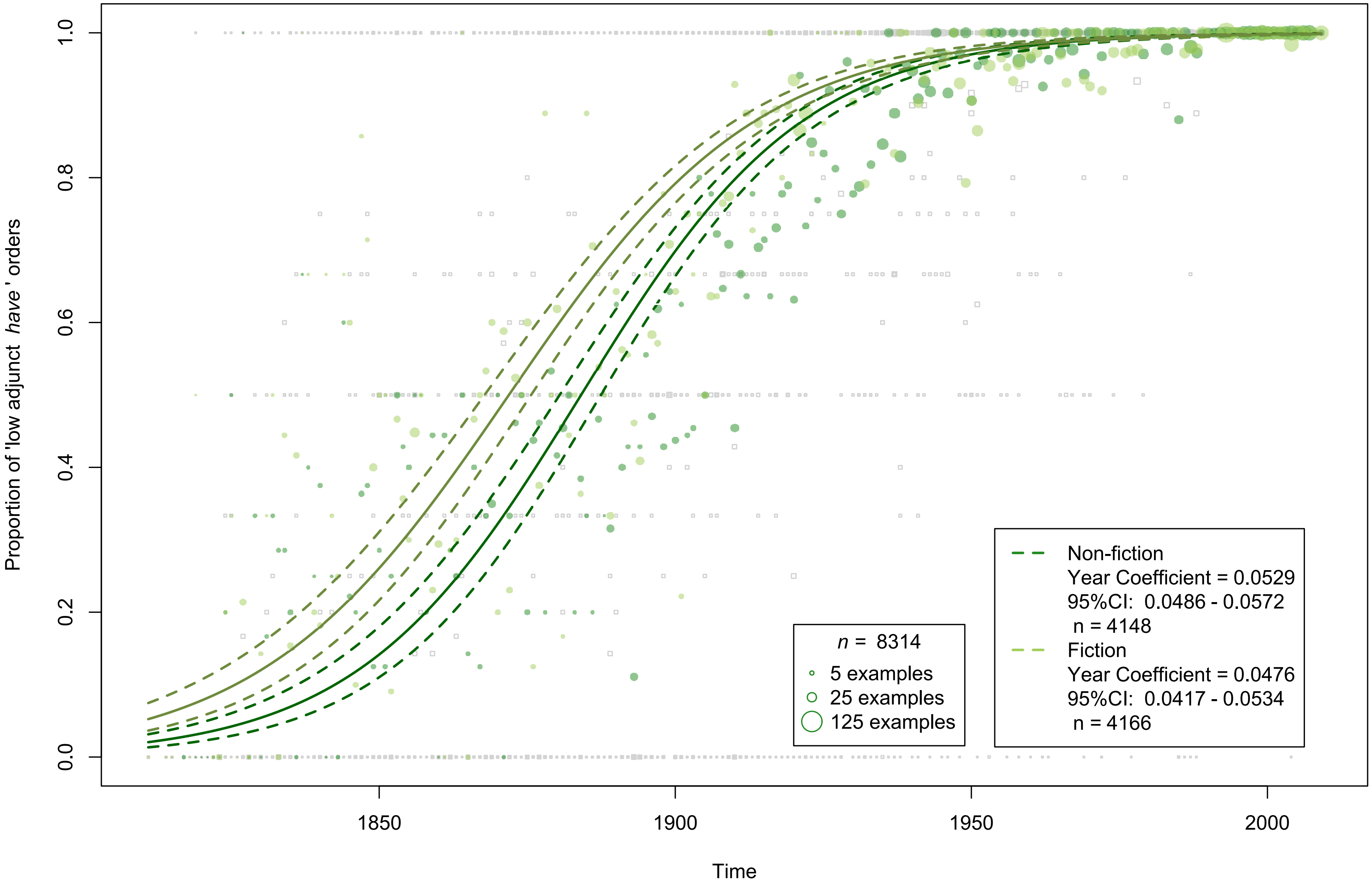



Figure 2.13: Illustration of the mixed-effects model for the adverb context

Why are the genre-related variables not significant in this model? In other words, why does the literary exploitation of conservative and innovative possessive *have* manifest in a significant difference between fictional and non-fictional writings for the negation and inversion data but not for the adverb context? Or put differently still, why does the (in)availability of *do*-support, but not the relative order between a VP-adjunct and possessive *have*, appear to function as a conspicuous marker of social identity in fiction? I have no definite answer to offer but would like to speculate on the question for a short moment.

First, it may be the case that word order variation is somehow less salient as a social marker than variability among functional elements. The socio-linguistic literature is replete with anecdotal references to the idea that purely syntactic variants do not lend themselves well to the acquisition of social meaning (for discussion see e.g., Hudson 1980, Winford 1996). For example, Hinskens follows a hierarchy of socially informative variability such that "the proportion of variable phenomena increases the closer one approaches the 'periphery' of the grammar, hence: syntax < morphology < phonology < phonetics" (1998: 160). Since one might qualify the relative order between low adjuncts and possessive *have* as 'syntax' and the presence or absence of the dummy auxiliary *do* as 'morphology' or 'morphosyntax,' this hierarchy would be in accordance with the observation that there is a weaker correlation between social dimensions and the use of the former than the use of the latter variable. On the other hand, various researchers have claimed that, in principle, all linguistic levels allow for socially meaningful variability to the same degree (e.g., Cheshire 2003).

Second, the conservative variant '*have* VP-adjunct' may be perceived as substantially less acceptable than conservative forms of negation or inversion in conjunction with possessive *have*. This is certainly plausible for the present day given that virtually all examples of possessive *have* sentences with low adjuncts have shown the innovative word order since the 1970s. As a consequence, authors may not be aware of the existence of the conservative word order option anymore, and hence never use it, but still associate direct negation or inversion with archaic connotations, and thus use these structures to communicate historical settings. The long tail of the regression curve may then have obscured previously held, conscious connotations associated with post-verbal VP-adjuncts. Indeed, if the above regression is re-run with data from before 1980 only, the difference in intercept and slope between fiction and non-fiction becomes marginally significant at the 5% level. Controlled, apparent-time grammaticality judgment tasks would likely be able to shed more light on the matter. For reasons of space, I will not consider the issue any further here, but move on to the fourth and final context that should reflect the underlying change in possessive *have*, verb phrase ellipsis.

### 2.3.6 Verb Phrase Ellipsis

This section deals with the data collection and analysis of verb phrase ellipsis (VPE) structures whose reconstructible meaning depends on an antecedent clause containing finite possessive *have*.

#### 2.3.6.1 Data Collection

The construction of practicable search queries for the retrieval of VPE structures proved substantially more difficult than for the previous contexts. As a syntactic phenomenon spanning multiple clausal units, VPE is very hard to find in an unparsed corpus like the COHA. The only feasible way to collect relevant data with the technology provided by the COHA infrastructure consisted in conducting relatively few search queries and filtering out a great many false positives manually. In fact, the vast majority of the retrieved hits turned out to be irrelevant for the purpose of this study. The strain of the manual correction process was thus particularly great for the VPE context, and human error may be a more serious source of noise in the data than before. Furthermore, the COHA search interface frequently malfunctioned under the pressure of the computationally intensive queries, indicating the wrong number of hits, producing large numbers of doublets, inexplicably returning seemingly irrelevant examples, and requiring several workarounds to download all the data. To make the task at least a little bit easier, I did not design any search queries for complex syntactic structures within VPE sentences, involving, for instance, adverbs, parentheticals or vocatives. Further, placeholders for subjects were exclusively operationalized as strings with a length of exactly one word in the form of the part-of-speech labels for pronouns and nouns. Finally, my search queries did not target contracted forms of possessive *have* in the source clause, such as *I've*. Instances of non-canonical VPE structures, complex subjects or contracted possessive *have* in the antecedent are thus only present in my data if they were accidentally picked up by my search queries.

The dependent variable could be realized in two ways. Either a source clause containing a finite form of possessive *have* preceded a target involving verb phrase ellipsis licensed after possessive *have* directly, '*have* $\Delta$,' or a target involving verb phrase ellipsis licensed by *do*-support, '*do* $\Delta$.'

All search queries made use of the COHA's collocation function. I required a non-contracted form of *have*, i.e., `have`, `has` or `had`, to occur up to 9 words to the left of the actual search term. This word functioned as a proxy for finite possessive *have* in a source clause. The actual search term, in contrast, was meant to retrieve the ellipsis structure in the target clause manifesting either as the conservative or as the innovative variant of the dependent variable. This is illustrated in (185) below.

(185) abstract search query structure:

... *have* ... { *have* $\Delta$ | *do* $\Delta$ }
(source) (target)

example of an intended hit:

My parents *have* lots of friends. At least I think they { *have* $\Delta$ | *do* $\Delta$ }!

The search queries for the VPE target clauses were designed around six independent variables. First, I coded, once again, for the inflectional form of the finite verb in the VPE clause. Examples were annotated as 'Present' for the searches involving the finite forms `have` and `do`, as 'Present 3rd singular' for `has` and `does` and as 'Past' for `had` and `did`.

Second, I traced the part-of-speech label of the word meant to instantiate the subject. If an example was retrieved by a search query with a pronoun, `[pp*]`, the subject type was coded as 'pronominal.' If the search query involved the part-of-speech label for a noun, `[n*]`, the resultant hits were labeled 'non-pronominal.'

Third, all VPE tokens were annotated for polarity. Search queries could either include a form of negation, `not` or `n't`, and were labeled as negative, or they did not involve such a form, and were then labeled as positive.

Fourth, an additional variable kept track of whether the negator, where present, was uncontracted, `not`, or contracted, `n't`. If the search query did not involve a negator, the contraction variable of the retrieved examples was not applicable, 'NA.' The contraction variable had ramifications for the relative order between the subject and other search query elements. Uncontracted negation in questions or tags forced the subject into a position in front of the negator, `[pp*]|[n*] not`, as in (186). Contracted negation, on the other hand, required the subject to come after negation, `n't [pp*]|[n*]`, as in (187).

(186) *'subject - negation'*

a. You **had** a daughter too, sir, **had you not** $\Delta$?
Deformed 1830

b. You **have** a brother in the Rebel army then, **do you not** $\Delta$?
SplitHeirs 1993

(187) *'negation - subject'*

a. Listen, the Biswangers **have** a pool, **haven't they** $\Delta$?
Mov:Swimmer 1968

b. They **have** submarines in Germany, **don't they** $\Delta$?
MotorBoysOnPacific 1909

In the same manner, the placement of the subject could vary for the search queries for positive polarity. Either the subject came before possessive *have*, `[pp*]|[n*] have`, as in (188), or inverted with it, `have [pp*]|[n*]`, as in (189). In either case, the 'NA' annotation remained in place. Hence, the contraction variable registered the difference between uncontracted and contracted negation, but not the difference between uninverted and inverted subject-verb orders.

(188) '*subject* have'

a. I **have** no military reputation, for if **I had** $\Delta$, it would doubtless be forever ruined
YaleRev 1913

b. I wish I **had** a skateboard. If **I did** $\Delta$, I'd join them.
Esquire 1999

(189) 'have *subject'*

a. but Dapple **has** the nature of a gentleman. So **have you** $\Delta$, Reuben
DayFate 1880

b. That's why I **have** life insurance, and so **do you** $\Delta$.
WashPost 1992

Fifth, I targeted two different kinds of sentence types. The final term in the search queries consisted of a set of punctuation signs representing the end of a clause. If this set did not involve a question mark, the sentence type was coded as 'declarative.' If the final set of punctuation signs did include a question mark, the variable took on the value 'interrogative.'

Finally, I manually assigned every example in the data set to a more specific clause type. The reason for this coding was that the two levels of the sentence type variable actually consist of many different constructions and may thus be somewhat uninformative or confusing on their own. Interrogative VPE structures comprise two distinct categories: (i) direct questions, which indicate a clear separation between source and target in terms of punctuation markers, different speakers or different grammatical subjects and (ii) tags, for which there is a union between source and target clauses as well as identity in grammatical subject reference, (190). Similarly, declarative VPE sentences comprise not only (iii) declarative matrix clauses, but also other phrases, namely (iv) argument, (v) adverbial, and (vi) comparative subordinate clauses. The type variable allows my data set to be surveyed easily for specific kinds of VPE ellipsis structures, say negative question tags, etc.

(190) *subcategories of 'interrogatives':*

a. *direct question*
"I **have** other duties to perform." "[**You have** $\Delta$]? What? To whom?"
LessonsInLifeAll 1885

b. *tag question*
I suppose [you **have** a car, too, Bernard, [**don't you** $\Delta$]]?
YetOtherWaters 1952

(191) *subcategories of 'declaratives':*

a. *matrix declarative clause*
"Then Europeans **have** no pigment?" "[Yes, **they have** $\Delta$], just like all other races of men;"
AdventuresYoung 1871

b. *argument subordinate clause*
... to find out whether antiparticles **have** antigravity. Some theorists think [that **they do** $\Delta$].
Time 1958

c. *adverbial subordinate clause*
the Reformed Episcopalians **have** a liturgy, [while **the Methodists have not** $\Delta$].
NYT-Ed 1876

d. *comparative subordinate clause*
Windows Media devices **have** less _ downloadable video available [than **iPods and PSPs do** $\Delta$]
PCWorld 2006

I conducted a total of 2 (dependent variable) * 3 (inflection) * 2 (subject type)* 2 (polarity) * 2 (contraction) * 2 (sentence type) = 96 search queries.[17] For instance, the query `have [pp*] ?` was designed to retrieve examples of interrogative, positive VPE clauses with pronominal subjects that invert with a present tense form of conservative *have*, whereas the query `[n*] did not .|;|,|!|-|/|:` targeted examples of declarative VPE structures with a non-contracted negator licensed by a past form of the innovative auxiliary *do*.

After all the data had been collected, I manually corrected precision errors. I removed all examples for which, somewhat surprisingly, the collocation term failed to produce an instance of the word *have*. Hence, there was no identifiable antecedent clause within the search space to the left of the ellipsis structure. Here and in the subsequent examples, I underline the string falsely matched by a particular search query.

(192) eyes twinkled at the prospect of a little fun; " no, we have n't. Now, boys, of course a good many
Trumps 1861

The second common kind of error consisted of hits for which the collocation term found some form of *have* other than possessive *have*. For example, the material preceding the search term may have involved an antecedent clause with a form of perfect *have*, (193a), causative *have*, (193b), or modal *have*, (193c).

[17] The search queries for verb phrase ellipsis are stored on the Dissertation DVD under `Chapter 2 - Possessive Have/06 Ellipsis Context/VP Ellipsis Queries.xlsx`.

(193) a. this heat? " " I have played baseball in worse than this and so have you , " she said. Then thinking
Harpers 1962

b. he did get off the subject. "And you had Pierre kill Powell, didn't you ," he went on in his merciless
Mignon 1862

c. man is backed up by society. He has to be obeyed. " " Has he ? If he forbade me to give alms – or
BlackLight 1930

I further deleted hits for which the collocation term itself had retrieved an instance of a VPE structure so that the specific function of *have* in the source could not be ascertained.

(194) a gesture: I have not! I have not! You know I have n't ! She continues: TITLE I can't afford
Mov:Greed 1925

Conversely, even if the collocation term did in fact pick up an instance of possessive *have*, the search term targeting the VPE structure frequently returned unsuitable structures. For instance, the placeholder intended for the subject often produced a possessed element instead, (195a), or the targeted ellipsis structure turned out to be completed after a parenthetical, (195b).

(195) a. for those who have homes to extend the blessing to those who have them not ,
Home 1835

b. I had five more to sacrifice for France!" "Do you , whoever you are, refuse to recognize your patriotic duty
Harpers 1917

I did not consider examples that showed a form of possessive *have got* or plain *got* in the antecedent clause.

(196) a. say. " " But you have got no fresh mutton now, have you ? " " Maybe Mr. Underhill has, " said
HillsShatemuc 1852

b. BIAGGI What? Someone tryin' to help a kid like you? JOSE You got no idea, do you? Somebody
Play:ManhattanTransits 1989

Furthermore, I deleted all hits that could not be grouped into any of the interrogative or declarative clause type categories illustrated in (190) and (191) above. Most importantly, I discarded all instances of VPE constructions in relative clauses. The reason for this decision was the following. There was no systematic way to differentiate between pertinent cases of ellipsis and irrelevant cases of object relativization for the conservative variant of possessive *have* in such structures. The ambiguity is informally sketched in (197a). The innovative variant, in contrast, would always have been unambiguous in this respect, as illustrated in (197b).

(197) a. New York, with one hundred and fourteen times the population of Nevada, has simply the same vote that Nevada has ,
i. relative clause with verb phrase ellipsis: **has** simply the same vote that **Nevada has** $\Delta$
ii. relativization of object: has simply [the same vote] that Nevada has _

NorthAmRev 1877

b. if Croatia had the oil and lobby money that Kuwait does , would our policy towards them be the same?
i. only relative clause with verb phrase ellipsis: **had** the oil and lobby money that **Kuwait does** $\Delta$
SanFran 1992

Hence, the inclusion of relative clauses with VPE structures dependent on possessive *have* would have unfairly biased the data set towards the conservative variant.

Next, I removed all hits with antecedent clauses that independently contained information on the status of possessive *have*. The reasoning behind this choice was that the presence of unambiguously conservative or innovative possessive *have* in the source probably increased the probability of finding the same form in the target as well. Most importantly, possessive *have* could be negated, as in (198), or the antecedent could feature inversion, as in (199).

(198) *negation in source clause*

a. And you have n't a horn and neither have I, nor do you play.
DistantTrumpet 1960

b. Indians don't have the luxury of pretending. Ultimately, neither do we.
PlanetIndiaHow 2007

(199) *inversion in source clause*

a. Have women more time? and if they have, why should they spend it in this Sisyphus task?
AsWeWereSaying 1891

b. Do you have Batman number 497? And if you do, I'd be willing to trade you my first edition Dark Knight
Play:BrotherAndres 1992

I removed cases of emphatic *do* in the source for the same reason - they likely biased the VPE clause towards the innovative form, (200).

(200) You do have a car, don't you?
RageAngels 1980

While the majority of such examples did in fact seem to prime the realization of the VPE structure on the antecedent clause, there were also a small number of unambiguously asymmetric examples, for which the status of possessive *have* was clearly different in source and target. This observation offers particularly dramatic support for the idea that different, competing grammatical options can be represented within the mental grammar of a single individual speaker or author. Example (201a) shows *do*-supported subject-auxiliary inversion in the antecedent clause, but *have* as the licensor of verb phrase ellipsis. Example (201b) involves direct negation of possessive *have*, but the VPE structure is formed with *do*. Finally, example (201c) features non-finite possessive *have* after emphatic *do* in the source, but the conservative variant of possessive *have* in the target.

(201) a. "He wants her to cash a hundred thousand dollars' worth of bonds." "Does she have that much?" "She has n't , but she can get it.
MovingTarget 1949

b. Everyone was offering me half-interests in a revolution I had n't the slightest interest in. Or did I ?
AdventuresTheStainless 1972

c. She did have a very retentive memory, had n't she?
WalshGirls 1943

I removed all instances of non-finite possessive *have* in the antecedent, for instance after modal auxiliaries, (202a), or after the non-finite marker *to*, (202b). I did so in order to make the VPE data set consistent with those of the previous contexts.

(202) a. With the existent opportunities in sales, the field should have greater appeal than it has.
BroadcastManagement 1968

b. She was determined to have beautiful hair, and she did.
Town 1950

I manually corrected the sentence type variable where the crude operationalization in terms of a punctuation sign returned a declarative or interrogative sentence that actually better fitted the alternative category. Examples are shown below.

(203) *example retrieved as 'declarative'; corrected to 'interrogative' (tag)*

a. and I do not remember that you have any of this pattern – have you , Mademoiselle?
VillageInnTheAdventures 1843

b. You have a secret yearning for a leg-shackle, do you , her being an innocent and Merton's sister and all that?
ThenComesSeduction 2009

(204) *example retrieved as 'interrogative'; corrected to 'declarative' (comparative)*

a. How do I know that you have any more right here than I have ?
MrTuttFindsWay 1945

b. you expect me, [...], to believe you have greater insight into my personality than I do ?
Mov:Frances 1982

The most challenging aspect of the VPE correction was imposed by the contextual indeterminacy of the VPE resolution procedure. As explained in section 2.2.1.6, the identification of a relevant antecedent clause for a VPE target is a pragmatic act. I thus had to decide on the basis of linguistic context provided by the concordance results if the meaning of an ellipsis should in fact be reconstructed from a phrase headed by possessive *have* or from some other verb phrase. I kept hits if I felt that the VPE structure stood for possessive *have* but deleted them otherwise. In the example below, the VPE clause allowed for two different reconstructions, either *I don't have charge* or an emphatic *I don't know.* Here, the former interpretation seemed to make better sense because of the subsequent clause, *I'm not assigned to that part.* Therefore, this example was retained.

(205) I don't know who in each area has charge – I do n't ! I'm not assigned to that part. Maybe the medic
TomorrowsSon 1977

The VPE structure in the following example could also be interpreted in two ways, either as *they do not have a responsibility* or as *they do not respond.* In my opinion, the subsequent string, *and have no good reason for doing so*, made the latter option more plausible than the former. Thus, this example was deleted.

(206) They have a responsibility to respond. If they do not , and have no good reason for doing so, then
FranksCampaign 1864

There were a fair number of opaque cases for which the linguistic context did not provide sufficient disambiguating clues. For instance, the following VPE clause could either be reconstructed as *you do think I have a lot of faults* or as *you do have a lot of faults.* It seemed impossible to be reasonably sure which interpretation was intended. Hence, this example was not kept in my data set.

(207) like in me; you think I have lots of faults, you know you do. " " I suppose I do, in a way, " he acknowledged
JohnWardPreacher 1888

Such reliance on pragmatic intuitions may have introduced some bias into the data set. However, most examples seemed to me relatively straightforward to evaluate and thus the subjective influence on data selection may not be very strong.

After the removal of irrelevant examples, the VPE data set was left with a total of 2,332 observations.[18] In addition to several examples given above (e.g., (186) - (189)), the following near minimal pair illustrates the variation under discussion.

(208) a. *conservative variant:* 'have Δ'
WASHINGTON: If we only **had** Lee's seven thousand! But **we haven't** Δ. You may order the retreat at once, Colonel.
WashingtonCrossing 1932

b. *innovative variant:* 'do Δ'
"Not everyone **has** a garage. **We don't Δ.**" "We have no car." "Exactly."
FieldStream 2006

I will now analyze the development of the variation between direct and *do*-supported VPE structures after a possessive *have* antecedent clause.

[18] The verb phrase ellipsis data set can be found on the Dissertation DVD under `Chapter 2 - Possessive Have/06 Ellipsis Context/VP Ellipsis Data.xlsx`.

#### 2.3.6.2 Results

The VPE data was fitted to a mixed-effects logistic regression with the same structure as the mixed-effects models of the preceding three contexts, i.e., the regression equation predicts the form of the ellipsis licensor from time, genre and their interaction while controlling for the random effect of individual texts. Table 2.14 presents the result.

```
formula = Have ~  Year + Genre + Year:Genre + (1 | Text),
   family = binomial, data = verb phrase ellipsis

Fixed effects:
                          Estimate Std.Error z-value p
Intercept                 -104.3   2.210     -47.20  <0.001***
Year                      0.05345  0.001129  47.33   <0.001***
Genre(non-fic→fic)        19.58    2.325     8.42    <0.001***
Year:Genre(non-fic→fic)   -0.00990 0.001190  -8.32   <0.001***


Random effect:
Text, N=1,693
Variance of random intercepts:  0.488

Null deviance:        3231 on 2331 degrees of freedom
Residual deviance:    2086 on 2327 degrees of freedom
AIC: 2096
```

Table 2.14: Logistic Regression Model 7: mixed-effects model for verb phrase ellipsis structures

The model predicts an increase in the log-odds of *do*-supported possessive *have*-dependent VPE structures by 0.0535 per year for non-fictional texts. At this rate of change, the innovative variant would spread from 1% to 99% of use within 172 years. The corresponding slope for fictional texts is estimated somewhat lower at 0.0535 - 0.0099 = 0.0436, a rate at which the same rise in innovative VPE forms would require 211 years. The innovative variant is expected to break through the 99%-threshold of use in 2038 and 2051 for non-fiction and fiction respectively. The model thus predicts that the VPE change will be the last to go to completion among the four linguistic contexts considered here. The VPE environment's lower base probability of showing innovative structures should not be overestimated, however, as its transitional period largely coincides with the other contexts developing *do*-support, negation and inversion. The time variable is a significant predictor in the model. This confirms the existence of the hypothesized change in the realization of licensors for VPE dependent on possessive *have*. The difference in intercept and slope between fiction and non-fiction are statistically significant as well. The verb phrase ellipsis data comes from 1,693 texts overall. With a variance of only 0.488 in their random intercepts, individual texts show considerably less variability in the proportion of innovative possessive *have* than they do in the other three contexts. For instance, texts are estimated to show with a probability of 95% a proportion of the innovative form between [27.8% – 72.2%] at the mid-point of the change, where the overall probability of finding innovative *have* is 50%. The comparatively narrow range of the interval, 44.5%, indicates to me only moderate variation between texts.

The model performs in a satisfactory way. It fits significantly better than a null model (where the null model is an intercept only model) (Likelihood Ratio Test, $\chi^2$=1420.85, df=4, $p$<0.001). The model fits the data reasonably well (Pseudo-$R^2_{marginal}$= 0.554, Pseudo-$R^2_{conditional}$= 0.612). The model has a high classification accuracy (C-index=0.910, classification accuracy = 84.3% vs. baseline: 51.5%).

The graph in figure 2.14 serves as an illustration of the model. Regression lines, year averages and text boxes should be interpreted as before. The long left tail of the regression curves highlights the relative conservativeness of the VPE context. The regression line for non-fiction is considerably steeper than the line for fiction, which represents the statistically significant difference in their slope.

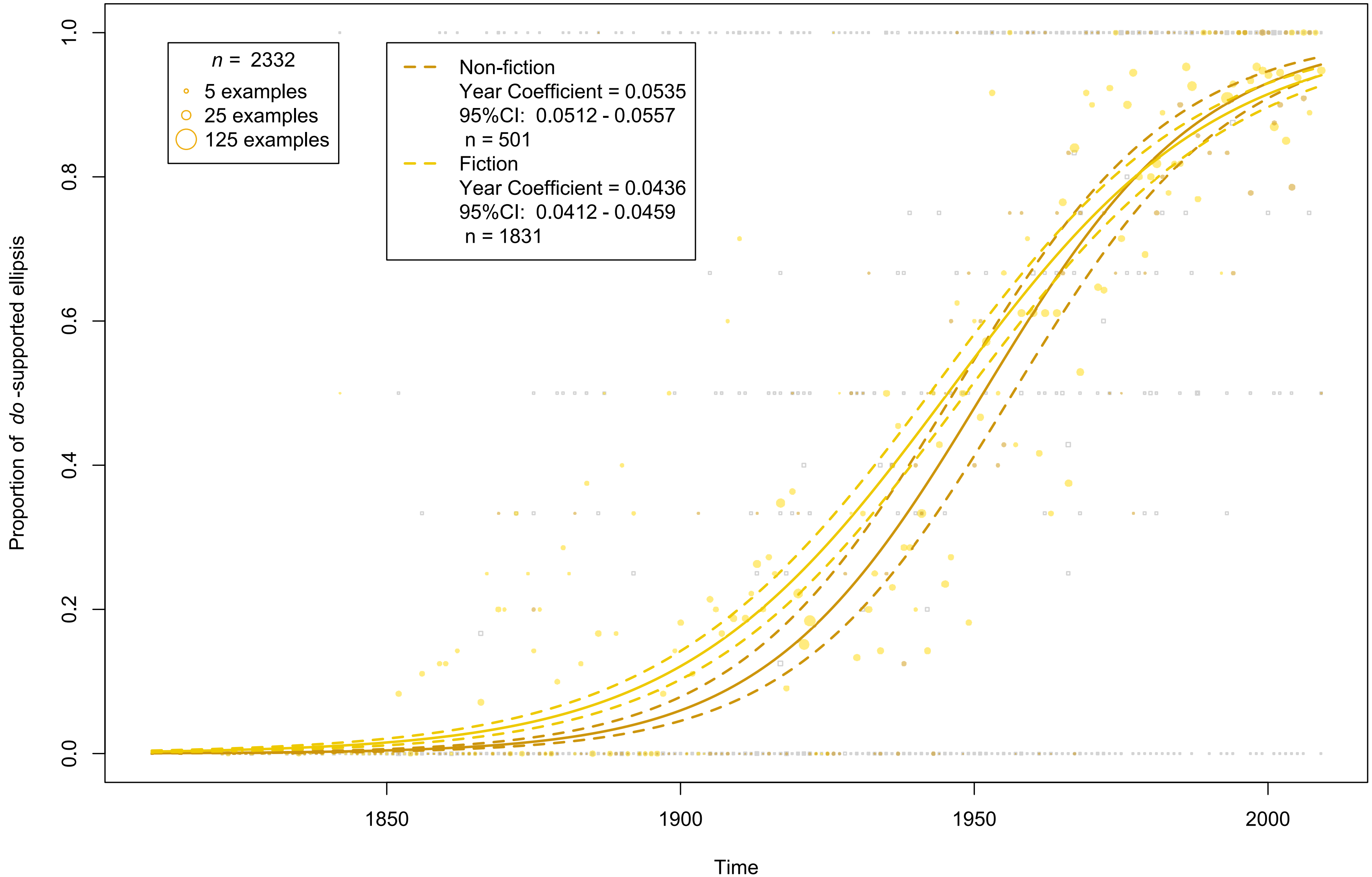



Figure 2.14: Illustration of the mixed-effects model for the verb phrase ellipsis context

The significant genre effect is likely to follow, once again, from the fact that fictional writings may contain forms of a variety that the COHA is not actually representative of. During the early stages of the change, fictional texts sometimes record vernacular uses of *do*-supported VPE structures with possessive *have* to portray lower-class speakers. A few examples are presented below.

(209) a. they **has** good times with me; and if **they don't** Δ, why, it's thar fault, and not mine.
UncleTomsCabin 1852

b. **Have** Ablisherners thar, **don't ye** Δ? them people that go agin the South?
AmongPines 1862

c. Well, chillen, I suppose them Kitterys **has** everythin' in real grander, **don't they** Δ?
OldtownFolks 1869

These excerpts show a fair number of non-standard features, such as number mismatches in subject-verb agreement (*they has*, *Kitterys has*), *them* as a demonstrative (*them people*, *them Kitterys*), and spelling indications of dialectal pronunciations of functional items (*everythin'*, *thar*) or lexemes (*ablisherners*, *chillen*). It is thus reasonable to assume that the *do*-supported possessive *have* ellipsis structures are a part of the depiction of idiomatic, vernacular speech. In contrast, fictional texts from the late stages of the change are likely to exploit the formal or archaic connotations of VPE clauses directly licensed by possessive *have* as a literary device to present a past age. Potential examples of this effect are shown below.

(210) a. "You **have** her permission?" "**I have** Δ." But I wondered how Sam Chamberlain would view my topic.
AmbroseBierceOne-eyed 2003

b. "You **have** musicians and jugglers at your villa, **have you not** Δ?" Bishop Freculf smiled.
NightBloomingNovel 2003

c. I mean to find out if it **has** any bearing on the murder. If **it hasn't** Δ, it's no affair of mine.
CutQuick 1993

Example (210a) is taken from Oakley Hall's book *Ambrose Bierce and the One-Eyed Jacks*, which fictionalizes the life of an actual historical figure, legendary newspaper person Ambrose Bierce (c. 1842 - 1914). The example in (210b) comes from Chelsea Quinn Yarbo's historical novel *Night Blooming*, whose story revolves around the medieval court of Charlemagne. Finally, example (210c) is uttered by Julian Kestrel, the protagonist of Kate Ross' detective story *Cut to the Quick*, set in England during the early 19$^{\text{th}}$ century Regency era. The authors may consciously have employed conservative possessive *have* VPE clauses in these excerpts to convey a sense of antiquated parlance in their period writing. The observed genre effect will then follow from these register fluctuations in the same way as explained earlier for the negation or inversion contexts.

This concludes my discussion of the change in possessive *have* in VPE constructions. Now that the data has been collected for the four, relevant linguistic contexts and independent regression models have confirmed that they manifest the change in possessive *have* in the predicted direction, I can move on to the last part of this chapter - an examination of the hypothesis that the rate of change is in fact identical in all of them.

## 2.4 Hypothesis Testing: Is there a Constant Rate Effect in the Possessive *have* Data?

One of the most widely cited formulations of the Constant Rate Hypothesis (CRH) states that, "when one grammatical option replaces another with which it is in competition across a set of linguistic contexts, the rate of replacement, properly measured, is the same in all of them." (Kroch 1989: 200). It is possible to apply this definition directly to the present case study. The set of competing grammatical options contains two members, innovative possessive *have* of category `V`, and the form it gradually replaces, conservative possessive *have* of category `I`. The set of linguistic contexts were deduced from a comparatively uncontentious theory of Modern English clause structure and comprise negation, inversion, adverb placement and verb phrase ellipsis constructions. I labored to measure the change in these contexts as properly as possible, collecting large samples to obtain precise statistical estimates, taking into account the effect of imitating styles by means of a genre variable, ruling out inadequate structures such as adverbs with a strong propensity for high adjunction, or correcting all precision errors in my data manually, etc. The rate of replacement can be operationalized straightforwardly as the coefficient of the time variable returned by a logistic regression fitted to the data. All that remains to be seen is whether or not the rates of change for the four linguistic contexts are in fact identical.

First, I will examine the ‘Year’ coefficients for each context returned by a simple logistic regression model that predicts the realization of the dependent, linguistic variable from time alone (for examples of this simple model structure, see logistic regression model 1 in table 2.3 or logistic regression model 5 in table 2.10). Here and in the following discussions, I will use the restricted set of VP-adjuncts described in table 2.12. The values can then be compared for identity or distinction on a more or less subjective basis. Table 2.15 presents the relevant pieces of information.

| Context | ‘Year’ coefficient | 95%-confidence interval |
|---|---|---|
| Negation | 0.04137 | [0.04054 – 0.04221] |
| Inversion | 0.03992 | [0.03850 – 0.04138] |
| VP Adjuncts | 0.03989 | [0.03797 – 0.04187] |
| VP Ellipsis | 0.04168 | [0.03846 – 0.04505] |

Table 2.15: Rates of change for the four contexts from independent, simple regression models

With values of 0.041, 0.040, 0.040 and 0.042 for negation, inversion, adverbs and ellipsis respectively, it is clear that the ‘Year’ coefficients are extremely similar to one another. The range between the highest and the lowest point estimates is very narrow at just 0.00179. This number would predict that the difference in the time it takes for the innovative form to rise from 1% - 99% of use varies by only 10 years among the four contexts. In the grand scheme of things, this difference is negligible. It thus seems quite plausible to conclude that all four contexts require the same amount of time to generalize the innovative variant of possessive *have*.

Secondly, one can investigate the CRH graphically by plotting the development of the innovative form in all contexts next to each other. Parallelism in the s-shaped curves would then indicate identity in the velocity of the change. However, it may be quite hard for the human eye to assess if s-shaped curves really appear aligned to each other. Such an evaluation is simpler if the regression lines show no curvature. Luckily, it is easy to transform the s-shaped probabilities modeled by the logistic function, (211), into straight-lined log-odds that follow a linear function through the logit transform, (212). The variable $\beta$ in the equations below stands for the ‘Year’ coefficient, i.e., the rate of change.

(211) $p = \dfrac{e^{a+\beta * year}}{1 + e^{a+\beta * year}}$ (s-shaped)

(212) $ln(\dfrac{p}{1-p}) = a + \beta * year$ (flat)

For reasons of comprehensibility, I shall give a few examples of the relation between proportions and log-odds. A proportion ($p$) of 50% corresponds to odds (*p/(1-p))* of 50% : 50%=1, and hence to log-odds (*ln(p/(1-p)))* of 0. Log-odds of 0 of an innovative form are therefore encountered at the midpoint of a change. Log-odds larger than 0 indicate percentages greater than 50% and vice versa. Thus, a proportion of, say, 30% is equal to odds of 30% : 70%=0.4286 and to log-odds of -0.8473 etc. Proportions of 100% and 0% correspond to odds of infinity and 0 and to log-odds of positive and negative infinity respectively. These values are therefore not usually included in graphs. The slopes of the flattened lines resulting from plotting not proportions but log-odds can then be appraised for identity in a straightforward way.

Figure 2.15 illustrates the development of possessive *have* in each linguistic contexts. The graph in the panel on the left side shows the average proportion of innovative *have* for every year as well as the s-shaped regression lines with their 95%-confidence intervals estimated from the simple logistic regression models just discussed. The graph in the panel on the right side represents the logit transform of the regression lines so that they appear flattened out. The data points and confidence intervals have been removed from this plot for greater clarity. The ‘Year’ coefficient for each context now corresponds directly to the slope of the linear functions.

As the graph shows, the lines look remarkably parallel to each other. The regression line for the inversion context seems perhaps slightly slanted against the line for the negation context. All in all, however, the inspection of the graphs suggest to me that the rates of change for the four linguistic contexts can reasonably be assumed to be identical.

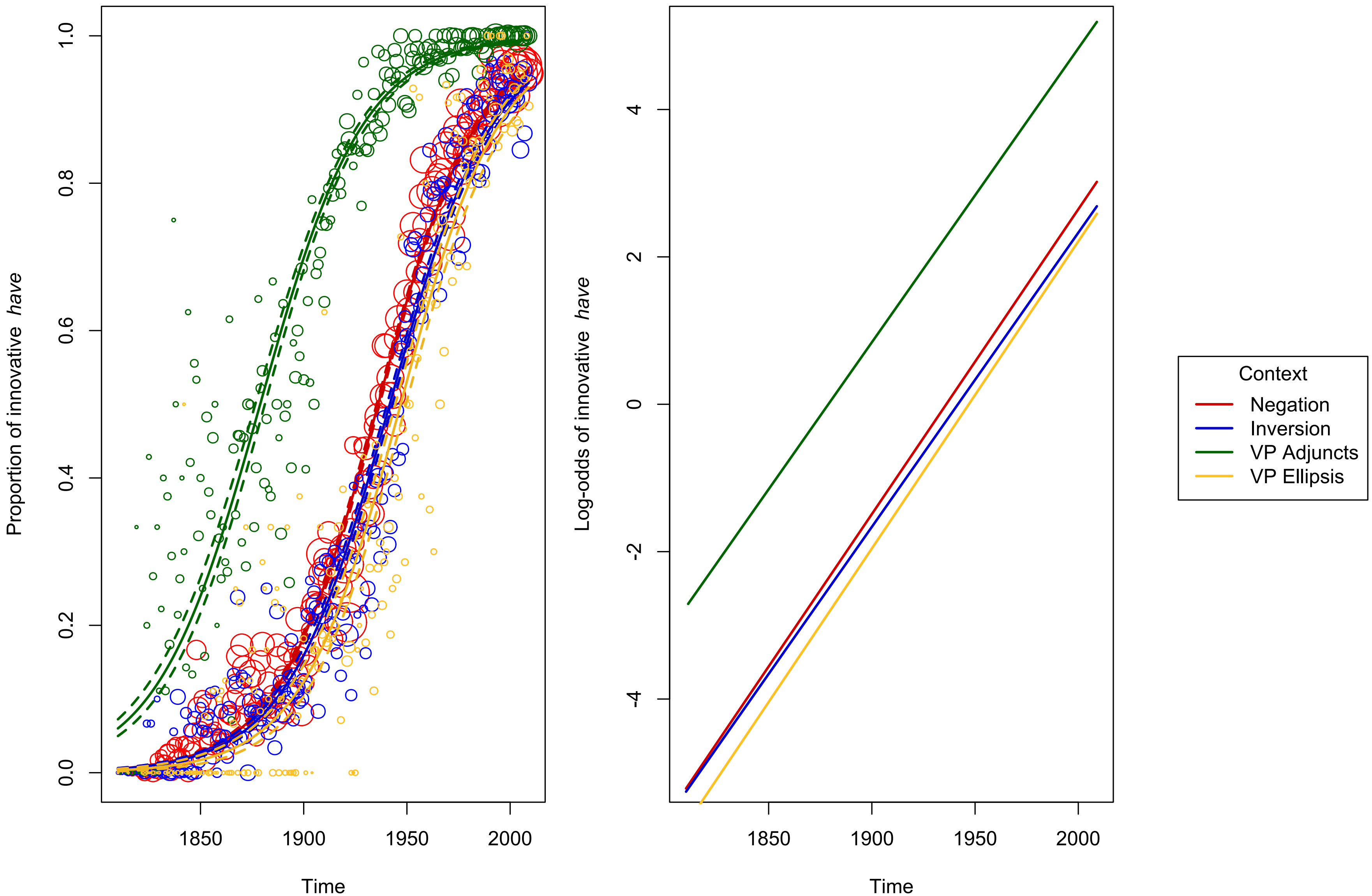



Figure 2.15: Graphical illustration of four independent, simple regression models for each context

Third, there are statistical tests to determine if there is sufficient evidence at some significance level to conclude that the rates of change differ between the four linguistic contexts. In other words, while it is not easily possible to directly check for the identity of the four time parameters, one can directly test if they are significantly different from one another. If they are not, this lends some credence to the correctness of the CRH.

In order to do this, two logistic regression models must be fitted to the combined data from all contexts. The joint data set has a sample size of 33,074 (negation) + 10,359 (inversion) + 8,314 (adverbs) + 2,332 (verb phrase ellipsis) = 54,079 observations. The first model will contain the crucial 'Year' variable as well as a 'Context' parameter that allows the model to keep track of the origin of each observation. The 'Context' variable can take on the values Negation, Inversion, VP Adjuncts or VP Ellipsis. In such a model, the four contexts may undergo the change at different overall probabilities of finding the innovative form, but they are all subject to the same global time effect, i.e., by necessity, the four contexts must change at the same rate. I will call this the 'reduced model.' The second model will also incorporate the 'Year' and 'Context' parameters, but will add an interaction term between those two variables. The interaction term will free up the slopes of the regression lines for each context so that they may change at different rates. I shall refer to this as the 'full model.' Now, if the four contexts develop independently of each other, the inclusion of this interaction term will significantly improve the model fit. In this case, one should reject the null hypothesis that there is no difference among the rates of change between the four contexts, and one would be justified in accepting the more complicated, full model over the simpler, reduced model. If, on the other hand, the four contexts change in the same manner, the addition of the interaction effect to a model with only main effects must not significantly improve the model fit. In that case, one should keep the null hypothesis of identity in the rates of change, and it would not be warranted to use the full over the reduced model. The second option would fail to refute the CRH and hence make it possible for it to be true.

Specifically, significance of the 'Year' - 'Context' interaction will be assessed by a likelihood ratio test. This test employs the difference in deviance between the full and reduced models as a test statistic. Informally speaking, a model's deviance (or residual deviance) is its goodness in terms of how likely the data are given its estimates for the specified parameters versus how likely the data are if a parameter is estimated for every single observation. Hence, a smaller deviance value indicates a better fit to the data. The deviance for the reduced model must necessarily be larger than the deviance for the full model since the additional parameter, here, the added interaction between time and context, will reduce deviance above the reduction brought about by fitting parameters for the main effects only. The question is whether this reduction in deviance is large enough to permit the inclusion of the extra regression parameter in the model. This can be tested by running a chi square test on the deviance difference with the number of degrees of freedom corresponding to the difference in degrees of freedom between the full and reduced models. Table 2.16 presents the result of this statistical test.

```
Model 1:  (reduced model)
  Have ~ Year + Context
Model 2:  (full model)
  Have ~ Year + Context + Year:Context
```

| Model | Df | Deviance | Difference Df | Difference Deviance | *p* |
|---|---|---|---|---|---|
| 1 | 54074 | 42174 | | | |
| 2 | 54071 | 42170 | 3 | 4.1973 | 0.2409 |

Table 2.16: Analysis of deviance table and likelihood ratio test for reduced and full simple models

The difference in deviance between the reduced and full simple models is 4.20. A chi square test can be run on this value with 3 degrees of freedom, one degree of freedom for every additional parameter in the full model, namely the changes in slope from one reference context, say Negation, to the other three contexts, i.e., the slope differences between Negation - Inversion, Negation - VP Adjuncts, and Negation - VP Ellipsis. This test yields a 24.1% chance of falsely assuming that the four contexts change at a different rate when in reality they change at the same rate, much higher than the customary 5% level. Thus, there is not sufficient evidence to conclude that possessive *have* changes at different rates in the four linguistic contexts. Even though the absence of evidence for a difference does not directly count as evidence for the absence of a difference, this finding does further add to the plausibility of the CRH.

Lastly, a series of likelihood ratio tests can be performed on reduced and full models structured as above so that each include data from only two contexts. This will test if there is reason to assume that any two contexts generalize innovative possessive *have* at significantly different rates from one another. The result of this pairwise comparison of contexts is presented in table 2.17. It shows for every context pair the difference in deviance between full and reduced model as a chi square value, the degrees of freedom, which is always 1 for the additional interaction term in the full model, and the corresponding $p$-value, the probability of committing a type 1 error.

| second context included in the model | first context included in the model: Negation | Inversion | VP Adjuncts | VP Ellipsis |
|---|---|---|---|---|
| Negation | - | - | - | - |
| Inversion | $\chi^2$=2.875, df=1, $p$=0.08996 | - | - | - |
| VP Adjuncts | $\chi^2$=1.8301, df=1, $p$=0.17612 | $\chi^2$=0.0006, df=1, $p$=0.98048 | - | - |
| VP Ellipsis | $\chi^2$=0.0326, df=1, $p$=0.8567 | $\chi^2$=0.9342, df=1, $p$=0.33377 | $\chi^2$=0.8479, df=1, $p$=0.35715 | - |

Table 2.17: Likelihood ratio tests for context-by-context comparisons, simple models

None of the models improve significantly by including the interaction term between the time and the contexts under investigation. Hence, there is no evidence to suggest that any two contexts develop innovative possessive *have* at different rates. The lowest $p$-value is obtained by a comparison between negation and inversion, which accords well with the slight deviation between those contexts observed during the visual inspection of figure 2.15 above, but it still exceeds the conventional 5% significance level. Again, one is thus not justified in rejecting the null hypothesis that the rates of change are identical between the contexts.

I shall now repeat these four examinations of the CRH but take the estimates for the 'Year' coefficient from mixed-effects logistic regression models. These models include variables for time, genre, their interaction, as well as varying intercepts for individual texts, as discussed extensively in the previous sections. I will thus examine if the data can still plausibly be assumed to display a constant rate effect when there is some control for the influence of register fluctuations and for random variation among texts.

First, I will compare the 'Year' coefficients returned by such models in a relatively subjective way. Table 2.18 lists the relevant values for each context measured in non-fictional and fictional texts. For more details on these estimates, see the illustrations of their models in figures 2.9 for negation, 2.10 for inversion, 2.13 for adverb placement, and 2.14 for verb phrase ellipsis.

| | Context | 'Year' coefficient | 95%-confidence interval |
|---|---|---|---|
| **Non-fiction:** | Negation | 0.05729 | [0.05578 – 0.05880] |
| | Inversion | 0.05976 | [0.05763 – 0.06189] |
| | VP Adjuncts | 0.05287 | [0.04859 – 0.05715] |
| | VP Ellipsis | 0.05345 | [0.05123 – 0.05567] |
| **Fiction:** | Negation | 0.04567 | [0.04409 – 0.04725] |
| | Inversion | 0.04554 | [0.04343 – 0.04766] |
| | VP Adjuncts | 0.04755 | [0.04173 – 0.05338] |
| | VP Ellipsis | 0.04355 | [0.04121 – 0.04589] |

Table 2.18: Rates of change for the four contexts from independent, mixed-effects regression models

The rates of change for the negation, inversion, adverb and ellipsis contexts, in that order, are 0.057, 0.060, 0.053 and 0.053 in non-fictional, and 0.046, 0.046, 0.048 and 0.044, in fictional texts, respectively. As before, these values strike me as very close to each other. The range between the highest and the lowest point estimate is 0.00689 implying a difference in the time needed for the change to go to completion of 21 years for the former and 0.004001 leading to a time variation of 18 years for the latter genre category. Again, the rather small divergence of these numbers suggests that the implementation of the change in possessive *have* essentially requires the same amount of time in each of the four linguistic contexts.

Secondly, a graphical evaluation of the CRH is afforded by figure 2.16. It presents the development of possessive *have* in the four contexts separately for the two genre categories on a probability and log-odds scale.

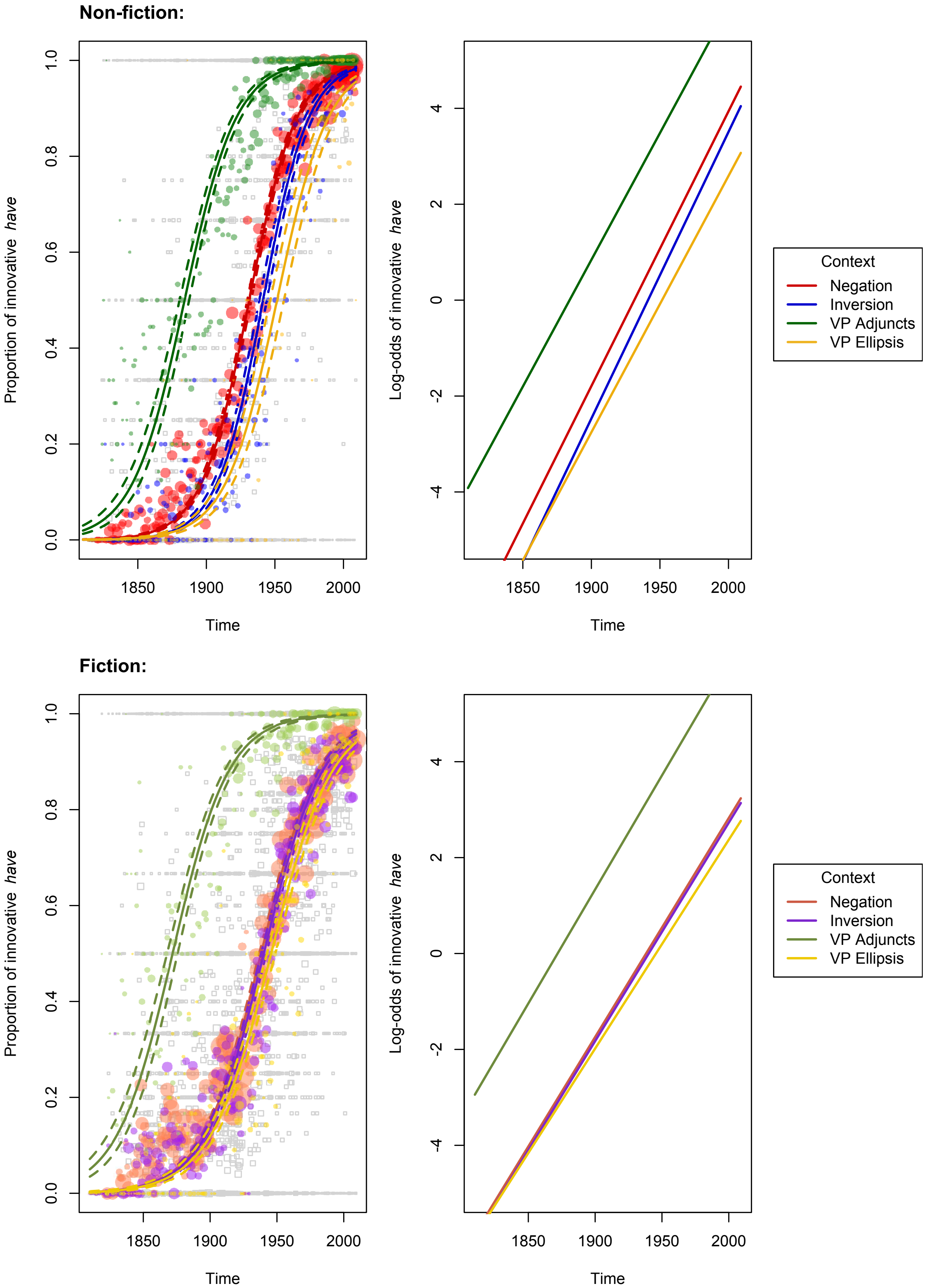



Figure 2.16: Graphical illustration of four independent, mixed-effects regression models for each context

The regression lines seem to run relatively parallel to each other. The adverb line may be somewhat flatter than the other three contexts when it is measured in non-fictional texts, and inversely slightly steeper when its slope is obtained from fictional texts. The rate of change for inversion in non-fictional writings appears to be marginally greater than those of the other contexts. Furthermore, the regression lines for negation, inversion and verb phrase ellipsis are essentially on top of each other for fiction whereas their baseline probabilities seem to be more divergent in non-fiction. Despite these minor discrepancies, the general parallelism between the regression lines leads me to affirm, once again, the presence of a constant rate effect in the data.

Third, one can test for a significant difference between the rates of change among the four contexts. Table 2.19 presents an analysis of deviance table comparing a full, mixed-effects model with, to a reduced, mixed-effects model without, a term for the interaction between the 'Year' and 'Context' variables. It also shows a likelihood ratio test on the difference in deviance between these two models.

```
Model 1:  (reduced model)
 Have ~ Year + Genre + Context + Year:Genre + (1 | Text)
Model 2:  (full model)
 Have ~ Year + Genre + Context + Year:Genre + Year:Context + (1 | Text)
```

| Model | Df | Deviance | Difference Df | Difference Deviance | $p$ |
|---|---|---|---|---|---|
| 1 | 54071 | 40021 | | | |
| 2 | 54068 | 40020 | 3 | 0.5817 | 0.9006 |

Table 2.19: Analysis of deviance table and likelihood ratio test for reduced and full mixed-effects models

As expected, the inclusion of the interaction term does not significantly improve the model fit. Hence, there is no justification for the assumption that the rates of change should be regarded as different for the four linguistic contexts. In fact, controlling for genre differences and variability among texts considerably strengthens the case for the CRH. The probability of committing an error in claiming that the four contexts develop at distinct rates is now a staggering 90.1%. While the test cannot technically demonstrate that the rates of change are identical, this result speaks as strongly in favor of the CRH as one could probably hope for.

Finally, I will test for significant differences between the rates of change between any two contexts. Table 2.20 presents the results of likelihood ratio tests for full and reduced, mixed-effects models that each include data from only two of the four linguistic contexts. As before, it reports the deviance difference as a chi square value, and its associated $p$-value on 1 degree of freedom.

| second context included in the model | first context included in the model: Negation | Inversion | VP Adjuncts | VP Ellipsis |
|---|---|---|---|---|
| Negation | - | - | - | - |
| Inversion | $\chi^2$=0.1781 df=1, $p$=0.67305 | - | - | - |
| VP Adjuncts | $\chi^2$=3.3584, df=1, $p$=0.06686 | $\chi^2$=1.7657, df=1, $p$=0.18392 | - | - |
| VP Ellipsis | $\chi^2$=0.4249, df=1, $p$=0.51448 | $\chi^2$=0.8822, df=1, $p$=0.34761 | $\chi^2$=2.4587, df=1, $p$=0.11687 | - |

Table 2.20: Likelihood ratio tests for context-by-context comparisons, mixed-effects models

Again, none of the differences in slope between any two contexts reach statistical significance at the customary 5% level. Thus, the data does not justify the belief that any two contexts develop innovative possessive *have* at rates different from one another. Instead, one must continue to operate under the null hypothesis, i.e., the assumption that there is no difference in the rates of change between different linguistic contexts. The lowest $p$-values are obtained for the differences between VP-adjuncts and the remaining contexts. I also noticed slight divergences between the slopes for low adjuncts and the other three contexts during the graphical evaluation of figure 2.16 above. The reason for this finding may be that the adverbs used for the measurement of the change have a high but not absolute chance of indicating the VP boundary - they may still occur in a high-adjoined position around 5% of the time. Uninformative parses of this kind may then cause a slight misestimation of the rate of change, but the distortion is not serious enough to induce a significant difference in the slope with respect to the other contexts and hence to refute the CRH.

The conclusion of this section is unequivocal: There is very good reason to assume that there is underlying identity in the 'Year' coefficients returned by logistic regression models that characterize the change in late Modern American English possessive *have* in negation, inversion, low adjunct and verb phrase ellipsis structures. The answer to the main question to be investigated in this chapter (see hypothesis in (130)) is thus affirmative – the linguistic innovation investigated in this case study does in fact seem to display a constant rate effect. By extension, this finding also lends additional support to the correctness of the Constant Rate Hypothesis as such. It is plausible to assume that linguistic changes with competing grammatical options feature identical rates of change across all contexts of use not as an exception, but as a generality.

## 2.5 Summary and Outlook

The rise of innovative possessive *have* in late Modern American English constitutes a straightforward, textbook example of a syntactic change. It displays with astonishing clarity all the characteristics of a common replacement change, as defined in chapter 1: It proceeds via competition between two, easily identifiable, theoretically well understood variants of the dependent linguistic variable. It completely replaces the conservative with the innovative form. It can be traced accurately in a large body of texts as there are no irreparable contaminations from prescriptive pressures or other cultural, scribal or historical confounding factors. Furthermore, the measurement of the change was conducted in a very careful manner and has produced one of the largest richly structured and manually corrected data sets discussed in the literature. For example, it exceeds the sample size of Ellegård's (1953) comparable data for the relevant contexts by a factor of more than 5 and the number of data points by a factor of almost 20. The substantial size of the data set permitted a very precise quantification of the rate of change. The detection of a constant rate effect with respect to four soundly motivated linguistic contexts thus means that the Constant Rate Hypothesis survived a very harsh falsification attempt in this case study. It would certainly have been possible for the slopes of any two different contexts to be significantly different from one another. The corroboration of a hypothesis should be commensurate to the unexpectedness of the correct predictions it makes.

> "A theory is tested not merely by applying it, or by trying it out, but by applying it to very special cases - cases for which it yields results different from those we should have expected without that theory, or in the light of other theories. In other words we try to select for our tests those crucial cases in which we should expect the theory to fail if it is not true. [...] It is an attempt to refute it; and if it does not succeed in refuting the theory in question - if, rather, the theory is successful with its unexpected prediction - then we say that it is corroborated by the experiment. **It is the better corroborated the less expected, or the less probable, the result of the experiment has been** [emphasis mine]." (Popper 1963: 150)

The affirmative answer to the research question of this chapter should therefore lead to a very substantial strengthening of one's confidence in the validity of the Constant Rate Hypothesis. In fact, the Constant Rate Hypothesis must be regarded as one of the best corroborated findings in all of quantitative historical linguistics.

The investigation of the change in American English possessive *have* leads to many open questions and possible future extensions of this research. First, the data sets created are rich enough to warrant further inspection. One could, for instance, analyze the influence of several variables that were registered as a side effect of the data collection process, such as different inflections or object elements. Most importantly, perhaps, one could explore the inversion data along the lines of Ellegård's (1953) division into negative, affirmative yes-no and *wh*-questions.

Secondly, it may be possible to test the CRH in additional linguistic contexts that should undergo the same change. For instance, possessive *have* should change its surface realization in negative imperatives (*Don't have too many plans!*) or with VP-topicalization (*She said we would have a discussion and have a discussion we did*). However, I suspect that such constructions are extremely rare and very difficult to extract from unparsed corpora with good accuracy. Similarly, possessive *have* should develop *do*-support for emphatic polarity, but it seems almost impossible to determine stress on conservative *have* in written data and hence to delineate the correct envelop of variation. Modal *have to* appears to innovate *do*-support in negation, inversion and VP-ellipsis structures as well (see e.g., Mair 2014 for many pertinent observations). A near minimal pair is shown below.

(213) a. *conservative variant: direct negation*
Here I **haven't to bow** and cringe to gentlemen of the aristocracy
EnglishwomanIn 1856

b. *innovative variant:* do-*supported negation*
I **don't have to bow down** to any one here.
PoorWhite 1920

Since modal *have* involves completely different subcategorization requirements than possessive *have*, the former would not strictly speaking be predicted to change at the same rate as the latter. If it does, this may entail interesting consequences for theories of the structure of the lexicon and polysemy. I did not analyze relative clauses with verb phrase ellipsis dependent on possessive *have* here, but they, too, are developing *do*-support at the moment but at a lower base probability, and may therefore offer a testing ground for grammaticality judgment experiments.

Thirdly, one could attempt to identify additional determinants of the choice between conservative and innovative possessive *have*. For instance, it has been claimed that the 'abstractness' of the possessed element (e.g., Butters 2001), the inalienability of possession, grammatical aspect (e.g., Traugott 1972) or the dynamicity - stativity distinction (e.g., Trudgill et al. 2002) may influence the realization of possessive *have* questions. Confirmation of the relevance of such factors in a large text sample could be illuminating for the question of how large base probability differences between different contexts would have to be in order to be recognized by speakers, or at least by professional linguists. Failure to demonstrate the importance of such factors might suggest that spurious claims about semantic distinctions between grammatical options that seemingly are in free variation actually reflect linguistic changes in progress. Closer examination of the social and geographical stratification of the diachronic variation may open up, among others, avenues into studying the divergence of English dialects (e.g., Varela Pérez 2007).

Fourth, it may be possible to run more elaborate statistical tests on the data sets I constructed. Formalisms are being developed that attempt to constrain constant rate effects more carefully (e.g., Kauhanen and Walkden 2016) and a large data set like the one used here might be helpful in designing some of their foundational experiments. It may also be feasible to conduct a special kind of power analysis on my data to estimate the probability of rejecting the null hypothesis when the alternative hypothesis is in fact true, which may conform more closely to a formal rendering of the Constant Rate Hypothesis than significance tests for type 1 errors.

Perhaps the most conspicuous open question concerns the cause of the change. Why did possessive *have* change its syntactic realization in the way it did? As outlined in chapter 1, *why*-questions regarding historical contingencies are notoriously difficult to answer. Nevertheless, I shall briefly share my thoughts on the matter in the following three arguments.

(1) In my opinion, the most promising kind of work towards a substantive understanding of general mechanisms of language change attempts to identify advantages of an intruding form over its conservative competitor. The most general identification of such fitness differences should be formulated in terms of better adaptability of the innovative variant within the structural system of the language as a whole. My idea for the present study is that *do*-supported possessive *have* structures may have higher parsing accuracy, which might lead to higher copying fidelity. Put differently, intended readings may be detected more sensitively with *do*-support signals, and every unambiguous parse of the new form would entrench it further in the mind.

Three potential examples are the following: First, 19$^{th}$ century speakers of American English could place only moderate confidence in a string such as *had not your* instantiating a negative declarative possessive *have* clause. The structure could also be resolved as a verb-first conditional, a negative question or a perfect with an unusual parenthetical. In contrast, 21$^{st}$ century speakers of American English can be almost entirely certain that the alternative string *did not have your* must communicate a negative declarative possessive *have* construction. Alternative parses are extremely unlikely. This hypothetical reasoning is supported by the fact that the false hits for the negation context overwhelmingly occurred with the conservative variant of possessive *have* (see (133) - (134)). On the other hand, this conjecture may become untenable once more syntactic context is taken into account. For instance, if a subject `DP` occurs in initial position, conservative variants cannot be interpreted as a verb-first conditional or negative question.

Second, a string such as '`DP` have not' could likely turn out to instantiate either a possessive *have* structure or a perfect in the conservative period. The same string cannot be followed by an object in the innovative period anymore. The fact that the innovative variant needs to anticipate only one likely continuation of such a string, namely with a past participle, whereas its conservative competitor had to delay certainty of the intended parse until either a past participle or an object was processed could possibly be an advantage that the former option

had over the latter.[19] However, it is not clear in how far anticipation plays a role in processing efficiencies. Third, strings such as *how many X have Y* are structurally ambiguous under the conservative variant of possessive *have*. They can be analyzed as both an object and a subject question. In contrast, the rise of *do*-support removes this ambiguity; *how many X do Y have* can now only be an object, *how many X have Y* only a subject question (see (151)).
Such recognizability advantages of the innovative variant must be weighed against analogous ones of the conservative form. For instance, the conservative form has the potential to disambiguate verb phrase ellipsis structures dependent on possessive *have* whose object contains a predicate. If a 19th century speaker continued the antecedent *A was afraid that B had something grave to announce* with the ellipsis structure *and he did* $\Delta$, its meaning could only be reconstructed as $\Delta$ = 'B did announce something grave,' whereas the ellipsis *and he had* $\Delta$ could only be interpreted as $\Delta$ = 'B did have something serious to announce.' A 21st century speaker will solely produce the *do*-support structure, but it now allows for the derivation of both meanings. If the parsing advantages of innovative possessive *have* were greater than those of its conservative equivalent, this would constitute a sufficient explanation for its rise.
Even if an explanation along these lines were correct, and even if it could be tested somehow, it would still be unclear why the change occurred at the time that it did. If recognizability or processing efficiency played a role, the change could have been triggered, in principle, at any point during the 16th to 19th centuries.
There might also be other contenders for specific, maybe even unique, kinds of advantages assigned to the new variant of possessive *have*. For example, it has been suggested that "there is a strong tendency to avoid overtly situating main verb *have* in Infl [=I], presumably because in Modern English [...] Infl is not a position from which Θ-roles are usually assigned." (Rohrbacher 1999: 196). An advantage might also manifest in socially motivated, possibly conscious, preferences for one form over another, i.e., socio-linguistic prestige. However, there is little evidence that either innovative or conservative possessive *have* ever carried consistent and firm value judgments. Rather, the features of the change are compatible with a change from below. However, some research suggests that similar forms of variations may be influenced by prestige, e.g., the variation between *have* and *have got* in Canadian English (Tagliamonte et al. 2010).

(2) I also find it appealing to invoke the traditional principle of analogy to elucidate the reason for the change even though its explanatory value may be debatable. By the time possessive *have* underwent its change, *do*-support had already solidly been implemented for main verbs. Thus, just as, say, weak past tense forms may be innovated on the basis of similar sounding irregular verbs, (214a), *do*-support in possessive *have* might be modeled after the majority rule for near-synonymous items, (214b).

(214) a. *drive* : *drove*
*dive* : X
X = *dove* (vs. standard *dived*)
b. *I own a car* : *I don't own a car*
*I have a car* : X
X = *I don't have a car* (vs. conservative *I haven't a car*)

(3) Finally, it has been suggested that the rise of *do*-support is intricately linked to, maybe even caused by, the loss of post-verbal adverbs (for discussion, not of a causal, but a very tight relation between the two phenomena, see e.g., Roberts 1985). The relation between the two developments is often discussed in relation with the 'Rich Agreement Hypothesis' (Vikner 1997). The matter is far too complex to be summarized here (for an overview, see e.g., Haeberli and Ihsane 2016), but it is noteworthy that my data is in fact compatible with some etiological connection between the change in adverb placement and the rise of *do*-support as the base probability of innovative constructions is appreciably higher for the former than for the latter contexts throughout the transitional period.

These considerations all seem potentially relevant and plausible to me. Still, they do not actually lead to testable predictions at the moment and must therefore be regarded as little more than mere speculations.

In many respects, this chapter provided a foundation for the subsequent discussions. First, it tried to justify the Constant Rate Hypothesis. In the next chapter, I will work under the assumption of its correctness, reasoning that the same processes that operate in a mental grammar to cause constant rate effects in the present also applied in the same way in the past. Rather than hypothesizing that plausibly related structures must undergo changes at the same rate, I will now presuppose the Constant Rate Hypothesis and argue that the demonstration of a constant rate effect among a set of linguistic contexts provides some evidence for their underlying relatedness.

[19] This idea was suggested to me by Eric Haeberli.

Secondly, the chapter also explained in some detail the statistical methods I use to analyze my quantitative data. I will employ the same techniques in the subsequent chapters, but without much further clarification. In particular, I will re-use the same kinds of plots and significance tests, as well as operationalizations of the rate of change in terms of the coefficient for the time variable returned by a logistic regression, and mixed-effects models that include random intercepts for different texts.

Last but not least, this chapter illustrated the general approach to the study of syntactic changes in this dissertation. I first built a formal grammar model, then derived from it predictions regarding the nature and direction of a change, and subsequently tested them on the basis of quantitative corpus data. It is true that the main hypothesis was actually very simple - different surface variations rest upon a mental grammar in which they are underlyingly related so that they should change at the same speed. But as a traditional English proverb says, ‘we must learn to walk before we can run.’ The following chapters will employ the same, general methodology but for more sophisticated research questions and more intricate changes in medieval English. The syntax of the variable patterns discussed will be less well-understood and the proposed grammar models will therefore be much more contentious. The quality of the data sets will be considerably poorer both as a result of smaller corpora and a greater number of factors corrupting the textual validity, such as translation effects, scribal intrusion during long manuscript transmissions or genres with greater distance to the spoken word. The hypotheses of the next chapter concern not just a lexical item’s change in syntactic category, but the introduction of new phrase structure rules as such, the interrelation between multiple changes, and the first theoretically predicted absence of a constant rate effect.

# 3

# Interaction between Multiple Syntactic Changes: The Loss of Old English Conjunct-Clauses[1]

*"Science has explored the microcosmos and the macrocosmos;*
*we have a good sense of the lay of the land.*
*The great unexplored frontier is complexity."*

– Heinz R. Pagels *The Dreams of Reason*, 1988

## 3.1 Introduction

The general research question of this chapter is the following: Is there a way to demonstrate interactions between multiple syntactic changes empirically? Specifically, is it possible to predict theoretically and measure quantitatively diverging trajectories of a primary syntactic change when it is or is not influenced by a secondary syntactic change that affects the grammar system at the same time?

The empirical material for this study is afforded by syntactic developments in Old and early Middle English (c. 850-1200 A.D.), referred to as 'Early English' in this chapter. In particular, my investigations focus on two primary structural changes of this period. Firstly, verb positions undergo substantial changes. Most notably, verb-final structures gradually disappear from the language. Secondly, there are modifications to the conditions on licensing argument preposing to clause-initial position, commonly called 'topicalization.' Overall, argument fronting becomes substantially less common. These two primary changes are affected by a secondary development that partially runs orthogonal to them. This third development concerns the loss of a grammatically important distinction between ordinary main clauses and those main clauses introduced by a conjunction like *and* 'and', *ac* 'but' or *ne* 'nor.' I will abbreviate ordinary main clauses as 'MCs' and henceforth refer to the conjunct main clauses as 'CCs.' The synchronic effects of this clause type difference on verb placement and topicalization can be described as follows.

[1]The idea for the syntactic analysis of Old English conjunctions presented in this chapter was first rudimentarily sketched in my Master's thesis (Zimmermann 2009). I have since considerably refined the analysis and presented my findings at the *Symposium on the History of English Syntax 11* in 2013, during an informal meeting of the *Treebanks* seminar at the University of Pennsylvania in 2014, at the *Lexical Functional Grammar Conference 19* in the same year, whose proceedings include preliminary statistical evidence for the synchronic Old English grammar model (Zimmermann 2014*b*), and at a *Séminaire de Recherche* in Geneva in 2015. This work profited greatly from the helpful feedback and encouragement by numerous participants of these conferences. I am also very grateful for instructive discussions with Ans van Kemenade, who kindly agreed to let me stay at Radbound University, Nijmegen for two weeks in August 2015 to provide me with feedback on several points of this research. Finally, I would like to explicitly thank the creators of the parsed corpora of historical English. The research described in this and the following chapter would have been impossible without their painstaking and conscientious annotation work.

(a) Verb placement. CCs are more commonly verb-final than MCs. That is to say, CCs place the finite verb after heavy non-subject material, as in (1a), substantially more frequently than MCs, which frequently show verb-second or specifically subject-verb order instead, as in (1b). In the sentences below and all subsequent Old English examples, the so-called 'Tironian note' '⁊' representing the word 'and' is rendered as the ampersand sign '&.'

(1) a. **&** [$_{\text{subject}}$ hie þa Romane] be þara biscepa lare hie swa cuce [$_{\text{verb}}$ *bebyrgdan*]
and they the Romans by the bishops' lore them thus alive buried
'And they, the Romans, buried them alive according to the bishops' teaching'
(coorosiu,Or_4:7.98.9.2019)

b. [$_{\text{subject}}$ Ic Ælfric munuc] [$_{\text{verb}}$ *awende*] þas boc of ledenum bocum to engliscum gereorde
I Aelfric monk turned this book from Latin books to English language
'I, monk Aelfric, translated this book from Latin books into English'
(coprefcath2,ÆCHom_II_[Pref]:.4)

I will not review the vast literature on the difference in verb placement between MCs and CCs here (for an excellent history of the identification of Old English conjunct clauses as a special clause type in this respect, see Bech 2015). I shall merely note that Old English scholar Bruce Mitchell seems to have had the most decisive influence on calling the difference to the attention of the wider research community. In his massive two-volume book *Old English Syntax*, he remarked factually that "conjunctions as *ond* and *ac* are frequently followed by the element order S ... V, which is basically subordinate" (Mitchell 1985: §1685) and criticized a large number of studies for their failure to take into consideration the important distinction between MCs and CCs (ibid: §1731). Subsequently, the majority of modern studies on Old English constituent order have added the value 'coordinate clauses' to the clause type variable, or have at least discussed the relevant difference (e.g., Koopman 1995; Pintzuk 1999; Bergen 2003; Ohkado 2004; Martín and García 2007; Ringe and Taylor 2014, to name but a few). Since conjunct clauses are now routinely treated as a special clause type, the differential verb placement pattern between MCs and CCs has become an extremely well-known fact in the field of Old English syntax.

A precise quantification of the strength of the clause type effect on verb positioning was missing from the literature for a long time. Instead, researchers made relatively unconstrained and imprecise generalizations to the effect that CCs "**tend** [emphasis mine] to be verb-final, like subordinate clauses" (Traugott 1992: 277), "**behave more like** [emphasis mine] subordinate clauses than main clauses" (Fuss and Trips 2002: 209), or "**often** [emphasis mine] have the verb-final word orders usually associated with subordinate clauses" (Fischer et al. 2000: 53) etc. Bech (2001) was the first scholar to provided reliable statistics. She correctly qualifies exaggerated claims regarding the preponderance of verb-final structures in CCs. Her counts show that, all in all, CCs are about three times as likely to place the verb late in the clause as MCs.

> [O]nly 15.3% (122 out of 795) of the OE conjunct clauses are verb-final. [...] We see that the word order of OE conjunct clauses varies greatly, and that the claim that they tend to be verb-final does not hold. What does seem to be the case, however, is that conjunct clauses are slightly more likely to be verb-final than non-conjunct clauses, of which only 5.4% have SXV word order [...]. (Bech 2001: 88)

(b) Topicalization. MCs more commonly involve a topicalized argument than CCs. That is to say, constituent fronting to a high, information-structurally prominent position is more likely in a sentence such as (2a), where the topic appears in clause-initial position, than in a sentence like (2b), which is introduced by a conjunction.

(2) a. [$_{\text{DP}}$ Todældu wæteru ] we lætað ut of urum eagum
divided waters we let out of our eyes
'We let the divided waters out of our eyes'
(cocura,CP:53.413.27.2858)

b. **and** [$_{\text{DP}}$ þone sang ] we sungon unseldon mid heom
and the song we sang frequently with them
'And we frequently sang the song with them'
(coaelive,ÆLS_[Swithun]:262.4384)

The distributional difference between MCs and CCs regarding topicalization is a lot less widely recognized than their divergence in terms of verb positioning. Kohonen (1978: 154) is, to the best of my knowledge, the only scholar who has quantitatively demonstrated a statistically significant effect of the difference between MCs and CCs on the presence of a topic. Using a relatively unconstrained measure of topicalization, he shows that MCs in Ælfric's *Catholic Homilies I* front a constituent to clause-initial position in 39% of all cases whereas *and/ac* clauses in this text do so only 7% of the time.

The distributional differences between MCs and CCs are unlikely to be due to Latin interference effects since they appear in both translated as well as original prose. For example, Mitchell (1964) and Bean (1983) discusses the clause type difference in the original Old English prose text *The Anglo-Saxon Chronicle* (see also Cichosz 2010, chapter 3). However, it may in fact be the case that Latin co-ordinating conjunctions influence constituent order in a comparable way (Devine and Stephens 2006: 163).

What is the reason for the curious clause type difference between MCs and CCs? It is my opinion that the basis for any explanation of a word order phenomenon should first be formulated in terms of a grammar model that provides a well-motivated phrase structural description of the pattern under investigation. Only subsequently should functional factors, such as information structure, other pragmatic constraints, or processing mechanisms be evoked as operations on these phrase structures as a possible reason for a speaker's choice to utter a particular variant from a set of truth-conditionally synonymous expressions.

The first step in the explanatory process, a formal description of Old English CCs, has, as far as I know, not been explicated beyond some rudimentary sketches. Campbell (1970) briefly entertains the idea that Old English coordinating conjunctions could be treated like subordinating conjunctions. Stockwell and Minkova (1990) mention (and reject) the same possibility. Haeberli (2001: 224-5, 2005: 8) allows `IP`s rather than `CP`s to be the input categories for coordination so that the left periphery of the clause would be truncated, which leads to the major distributional anomalies observed in CCs. Since a more satisfying formalization of CCs is still missing from the literature, I decided to focus my attention on a detailed computational grammar model outlining the internal structure of MCs and CCs.

The second step in the explanation, the identification of functional factors influencing the choice between different word order alternatives, has received a little more attention in the literature. For instance, Traugott considers a pragmatic explanation for verb-final CCs, hypothesizing that they "can be attributed to the fact that, from a discourse perspective, co-ordinate clause[s] elaborate on the initial main clause and in this sense modify it" (1992: 277). Bech bases her description on simple word order strings, such as *SXV*, *XVS* etc., and lets pragmatic factors exploit these patterns, concluding that "the pragmatic nature of SXV clauses, i.e., the way they must be organized in terms of information structure, implies that many of them will be conjunct clauses" (2001: 192). Pintzuk (1995) points out a priming effect on the occurrence of verb-final CCs. She shows that the probability of finding a verb-final CC is significantly elevated if the preceding MC is also verb-final. Explanations of this kind will not feature prominently in this chapter.

I would like to emphasize that I do not want to present my grammar model as a direct alternative to such functional factors. In fact, it is my hope that my model could in principle be married to accounts of information-structural word order determinants in order to explain a larger amount of the word order variation encountered in the Early English text data. I am also sympathetic to the idea that a change in extra-syntactic determinants of syntactic patterns can somehow have long-term effects leading to changes in the grammatical system itself. My grammar model will simply not deal with issues of usage because a description of linguistic structures takes precedence over a description of how they are used, and because functional factors (provided that they remain relatively stable) are not crucial in answering the larger research question of whether one can identify interactions between multiple syntactic changes in corpora.

This chapter is structured as follows. I will first describe the textual basis for the empirical generalizations reported here. In the next part, I will construct an LFG model of core phenomena of Old English clausal syntax. It will contain several pairs of phrase structure rules that I assume to be in competition in Early English such that one syntactic rule gradually drives out the other. To be precise, I propose changes in verb placement, topicalization and the category of conjunctions. I will then derive from the grammar model specific, quantitatively testable predictions regarding the synchronic distribution of word order patterns in MCs and CCs as well as diachronic interactions between changes in verb placement and the decline in topicalization with the changing syntactic status of conjunctions. In the subsequent part, I will examine the potency of my model by sequentially testing each of these hypotheses. In so doing, I will also address my research question as I will discover whether the presupposed interactions between the syntactic changes manifest in an empirically appreciable way or not. Finally, I will discuss my findings.

## 3.2 The Corpus

The Old English material for this chapter is taken from the syntactically parsed *York-Toronto-Helsinki Parsed Corpus of Old English Prose* (YCOE) (Taylor et al. 2003). All examples, their citations and the statistical counts I present rely on this database. Further, I follow the annotation scheme of the YCOE to determine the nature of a syntactic structure, in particular whether a clause is to be counted as main, coordinated, or subordinate.

There are a number of advantages that result from using parsed corpora like the YCOE. (i) Time savings. It could easily take months or even years to filter out particular constructions from millions of words of running text manually. With a parsed corpus, such a task can be completed in a few moments. (ii) Large numbers. The YCOE contains almost the entire corpus of surviving Old English prose texts. The statistical key figures derived from it are therefore likely as precise as they will ever become. (iii) Reduced human error. Humans can easily overlook examples or tally them up inconsistently. Collecting data from parsed corpora with automated search queries, in contrast, causes few mistakes because computers always count correctly. (iv) Objectivity. The inclusion or exclusion of a particular example in a group depends on explicit search criteria in a query file and is thus independent of the researcher producing them. (v) Reliability. The numbers obtained can be reproduced and double-checked with relative ease. (vi) Construct validity. Researchers often use concepts, such as ‘subject,’ that are defined only vaguely. When working with a syntactic corpus, one is automatically forced to operationalize syntactic notions on the basis of formal annotations, say nominative case. Thus, the definition will directly correspond to a construct one aims to measure.

However, reliance on parsed corpora also has certain downsides. The first of those is relatively subjective. During my work with the YCOE, I often experience a certain distance from the data. I do not have to read the examples in context anymore, as they merely show up as isolated instances in the output files. I am thus not always confident that all retrieved examples truly belong to the pattern that I am interested in. It is also tempting to lump all hits together as if they were produced by one individual. In this way, research questions revolving around textual differences and the fine-grained intricacies of the Old English material - scribal errors, translation effects, style, narrative intent, dialectal differences etc. - fade into the background. Secondly, many syntactic structures are ambiguous, but the YCOE can only record one parsing solution. Even worse, the YCOE contains definite parsing errors. Hence, the increased objectivity comes at the expense of reduced accuracy. While all researchers can now work with the same data set, they will also all work with the same errors. Several annotation mistakes have been corrected over the years. However, I will use here the uncorrected, official release of the YCOE so that other researchers can replicate my findings more easily.

Perhaps the largest problem of the YCOE for my work is the uncertain chronology of the text files it contains. The YCOE follows the Helsinki dating system (Kytö 1996), which depends on manuscript dates to place a text into one of four centennial periods: OE1 before 850, OE2 850-950, OE3 950-1050, OE4 1050-1150. If a manuscript date is uncertain or a text’s presumed date of composition diverges substantially from the date of its manuscript witness, mixed periods are possible, such as OE2/4 etc. In my opinion, syntactic studies should privilege dates of composition over manuscript dates because it seems reasonable that syntactic structures are relatively resistant to modifications during the manuscript copying process. Furthermore, the logistic regression models I use to investigate linguistic changes require a particular year code for every text.

I therefore decided to assign to every text file in the YCOE an approximate year of composition. The dates are taken from a review of the secondary literature available for every text. Editions, journal articles and anthologies often report reasons and arguments for particular dates of composition. The dates are also based on some provisional statistical dating techniques relying on text-internal syntactic criteria (Zimmermann 2014*a*). Unfortunately, I cannot justify my dating choice for every single text here because of space and time constraints. It cannot be emphasized enough that the dating techniques I used are fraught with problems and must hence be regarded as only <u>approximate</u>, and that they must be revised as new evidence becomes available and thus must also be regarded as <u>provisional</u>. Nevertheless, the dates are not arbitrarily picked but chosen according to external standards and the dates of the largest, hence likely most influential, texts are actually relatively precise. I therefore believe my dates to be justifiable and accurate enough for the purposes of this study.

There are exactly 100 text files in the YCOE. Table 3.1 - 3.2 lists each of them, along with a more customary text name, and the approximate year of composition I chose. I ignored some text files because they are duplicates or otherwise inappropriate. Those files are indicated in italics. Some text files were merged into one. I indicate such amalgamations with curly brackets. Other texts had to be modified. Alterations I made to a file are indicated with an asterisk and described in more detail in the paragraphs below.

| Number | File Name | Text Name | Assigned Year |
|---|---|---|---|
| 1 | coadrian.o34 | Adrian and Ritheus | 1035 |
| 2 | coaelhom.o3 | Ælfric, Supplemental Homilies | 991 |
| 3 | coaelive.o3 | Ælfric's Lives of Saints | 997 |
| 4 | coalcuin | Alcuin De Virtutibus et Vitiis | 1005 |
| 5 | coalex.o23 | Alexander's Letter to Aristotle | 870 |
| 6 | coapollo.o3 | Apollonius of Tyre | 1025 |
| 7 | *coaugust* | *Augustine* | *not used, too small* |
| 8 | cobede.o2 | Bede's History of the English Church | 880 |
| 9 | cobenrul.o3 | Benedictine Rule | 965 |
| 10 | coblick.o23 | Blickling Homilies | 950 |
| 11 | coboeth.o2 | Boethius' Consolation of Philosophy | 897 |
| 12 | cobyrhtf.o3 | Byrhtferth's Manual | 1011 |
| 13 | cocanedgD | Canons of Edgar (D) | 1010 |
| 14 | cocanedgX | Canons of Edgar (X) | |
| 15 | cocathom1.o3 | Ælfric's Catholic Homilies I | 991 |
| 16 | cocathom2.o3 | Ælfric's Catholic Homilies II | 994 |
| 17 | cochad.o24 | Saint Chad | 865 |
| 18 | cochdrul | Rule of Chrodegang of Metz | 985 |
| 19 | cochristoph | Saint Christopher | 925 |
| 20 | cochronA.o23* | Anglo-Saxon Chronicle A* | duplicate content removed; |
| 21 | cochronC* | Anglo-Saxon Chronicle C* | remaining material grouped |
| 22 | cochronD* | Anglo-Saxon Chronicle D* | into four files; |
| 23 | cochronE.o34* | Anglo-Saxon Chronicle E* | 891, 964, 1060, 1105 |
| 24 | cocura.o2* | Cura Pastoralis* | 891 |
| 25 | cocuraC* | Cura Pastoralis (Cotton)* | |
| 26 | codicts.o34* | Distichs of Cato | 965 |
| 27 | codocu1.o1* | Documents 1 (O1)* | grouped into three periods |
| 28 | codocu2.o12 * | Documents 2 (O1/O2)* | 9th, 10th and 11th century charters |
| 29 | codocu2.o2* | Documents 2 (O2)* | average years of the three sets: |
| 30 | codocu3.o23* | Documents 3 (O2/O3)* | 843, 960, 1020 |
| 31 | codocu3.o3* | Documents 3 (O3)* | |
| 32 | codocu4.o24* | Documents 4 (O2/O4)* | |
| 33 | coeluc1 | Honorius of Autun, Elucidarium 1 | 1115 |
| 34 | coeluc2 | Honorius of Autun, Elucidarium 2 | |
| 35 | coepigen.o3 | Ælfric's Epilogue to Genesis | 1000 |
| 36 | coeuphr | Saint Euphrosyne | 985 |
| 37 | coeust | Saint Eustace and his Companions | 1030 |
| 38 | *coexodusP* | *Ælfric's translation of Exodus* | *not used; largely duplicate of no. 80* |
| 39 | *cogenesiC* | *Ælfric's Translation of Genesis* | *not used; largely duplicate of no. 80* |
| 40 | *cogregdC.o24* | *Gregory's Dialogues (C)* | *not used; seems too unreliable* |
| 41 | *cogregdH.o23* | *Gregory's Dialogues (H)* | *not used; seems too unreliable* |
| 42 | coherbar | Pseudo-Apuleius, Herbarium | 940 |
| 43 | *coinspolD.o34* | *Wulfstan's Institute of Polity (D)* | *not used; largely duplicate revision of no. 44* |
| 44 | coinspolX | Wulfstan's Institute of Polity (X) | 1010 |
| 45 | cojames | Saint James | 1030 |
| 46 | colacnu.o23 | Lacnunga | 1020 |
| 47 | colaece.o2* | Bald's Leechbook* | books 1 and 2: 890, book 3: 1010 |
| 48 | colaw1cn.o3 | Laws, Cnut I | 1020 |
| 49 | colaw2cn.o3 | Laws, Cnut II | 1020 |
| 50 | colaw5atr.o3 | Laws, Æthelred V | 1008 |

Table 3.1: Overview over the YCOE text files and their approximate dates of composition

| Number | File Name | Text Name | Assigned Year |
|---|---|---|---|
| 51 | colaw6atr.o3 | Laws, Æthelred VI | 1008 |
| 52 | colawaf.o2 | Laws, Alfred | 893 |
| 53 | colawafint.o2 | Alfred's Introduction to Laws | 893 |
| 54 | colawger.o34 | Laws, Gerefa | 1040 |
| 55 | colawine.ox2 | Laws, Ine | 871 |
| 56 | colawnorthu.o3 | Northymbra Preosta Lagu | 1025 |
| 57 | colawwllad.o4 | Laws, William I, Lad | 1110 |
| 58 | coleofri.o4 | The Vision of Leofric | 1060 |
| 59 | colsigef.o3 | Ælfric's Letter to Sigefyrth | 1005 |
| 60 | *colsigewB* | *Ælfric's Letter to Sigeweard (B)* | *not used; largely duplicate fragment of no. 61* |
| 61 | colsigewZ.o34 | Ælfric's Letter to Sigeweard (Z) | 1005 |
| 62 | colwgeat | Ælfric's Letter to Wulfgeat | 1000 |
| 63 | *colwsigeT* | *Ælfric's Letter to Wulfsige (T)* | *not used; largely duplicate fragment of no. 64* |
| 64 | colwsigeXa.o34 | Ælfric's Letter to Wulfsige (Xa) | 1000 |
| 65 | colwstan1.o3 | Ælfric's Letter to Wulfstan I | 1002 |
| 66 | colwstan2.o3 | Ælfric's Letter to Wulfstan II | 1010 |
| 67 | comargaC.o34 | Saint Margaret (C) | 1050 |
| 68 | comargaT | Saint Margaret (T) | 1030 |
| 69 | comart1 | Martyrology, I | 975 |
| 70 | comart2 | Martyrology, II | 920 |
| 71 | comart3.o23 | Martyrology, III | 900 |
| 72 | comarvel.o23 | Marvels of the East | 870 |
| 73 | comary | Mary of Egypt | 985 |
| 74 | coneot | Saint Neot | 1065 |
| 75 | conicodA | Gospel of Nicodemus (A) | 945 |
| 76 | *conicodC* | *Gospel of Nicodemus (C)* | *not used largely duplicate of no. 75* |
| 77 | conicodD | De Ascensione (Nicodemus D) | 1025 |
| 78 | *conicodE* | *De Ascensione II (Nicodemus E)* | *not used largely duplicate of no. 77* |
| 79 | coorosiu.o2 | Orosius | 898 |
| 80 | cootest.o3 | Heptateuch | 1000 |
| 81 | coprefcath1.o3 | Ælfric's Preface to Catholic Homilies I | 991 |
| 82 | coprefcath2.o3 | Ælfric's Preface to Catholic Homilies II | 994 |
| 83 | coprefcura.o2 | Preface to Cura Pastoralis | 891 |
| 84 | coprefgen.o3 | Ælfric's Preface to Genesis | 1000 |
| 85 | copreflives.o3 | Ælfric's Preface to Lives of Saints | 997 |
| 86 | coprefsolilo | Preface to Augustine's Soliloquies | 899 |
| 87 | coquadru.o23 | Pseudo-Apuleius, Medicina de Quadrupedibus | 940 |
| 88 | corood | History of the Holy Rood-Tree | 1030 |
| 89 | cosevensl | Seven Sleepers | 985 |
| 90 | cosolilo | St. Augustine's Soliloquies | 899 |
| 91 | cosolsat1.ox4 | Solomon and Saturn I | 1035 |
| 92 | cosolsat2 | Solomon and Saturn II, Pater Noster Dialogue | 870 |
| 93 | cotempo.o3 | Ælfric's De Temporibus Anni | 996 |
| 94 | coverhom | Vercelli Homilies* | Vercelli Homily 1 dated 910, the rest 950 |
| 95 | coverhomE | Vercelli Homilies (E) | 910 |
| 96 | coverhomL | Vercelli Homilies (L) | 980 |
| 97 | covinceB | Saint Vincent (Bodley 343) | 1010 |
| 98 | covinsal | Vindicta Salvatoris | 945 |
| 99 | cowsgosp.o3 | West-Saxon Gospels | 990 |
| 100 | cowulf.o34 | Wulfstan's Homilies | 1010 |

Table 3.2: Overview over the YCOE text files and their approximate dates of composition, cont.

(1) The *Anglo-Saxon Chronicle*, a collection of annals written in Old English from the $9^{th}$ century on, is contained in the YCOE in four different text files that are based on four different manuscripts. There is a lot of duplicate material shared between those four versions. I removed the duplicate content. Subsequently, I re-grouped the remaining annals into four text files with the assigned years 900, 964, 1060 and 1105. The periodization is based on annal numbers and the transitions from one copying scribe to another. I hope that the dates of the four files roughly corresponds to the time that the language they contain is representative of. However, more work is required to properly dissect this text.

(2) The base manuscript of *Cura Pastoralis*, Bodleian Hatton 20, is contained in the YCOE file cocura.o2.psd. Section 33 (tokens 1430-1499) is defective and replaced with the duplicate material from a second manuscript, Cotton Tiberius B.XI, contained in the YCOE file cocuraC.psd.

(3) Anglo-Saxon charters are among the most precisely datable Old English texts. I isolated individual charters contained in the six codocu files according to their Sawyer number, and re-grouped them into three periods. The first period contains the $9^{th}$ century charters 98, 126, 204, 218, 1188, 1195, 1197, 1200, 1283, 1445, 1482, 1500, 1510 and their average Sawyer date is 843. The second set comprises the $10^{th}$ century charters 1447, 1494, 1501, 1506, 1533 with an average year of 960. The final group covers the $11^{th}$ century charters 1394, 1399, 1454, 1458, 1460, 1467, 1473, 1486, 1487, 1503, 1530, 1534, 1471 and their average date of composition is 1020.

(4) The medical text *Bald's Leechbook* comprises three distinct books. The first two were likely translated as a part of King Alfred's educational reform program and have been assigned to the year 890; the third may have been composed later and was dated to the year 1010.

(5) The Vercelli Homilies were split into two files. The first contains only Homily number one, which was probably written earlier than the others. It was assigned to the year 910. The second file contains the rest of the text collection. It was dated to 950.

I also collected data from the earliest extant Middle English texts up to c. 1200. The corpora that these text files were taken from will be described in more detail in the following chapter. On account of the dearth of texts from the early Middle English period, I also included poetic records from this era. Table 3.3 summarizes the 15 early Middle English texts and their approximate date of composition I used in this study.

| Number | File Name | Text Name | Assigned Year |
|---|---|---|---|
| 1 | CMPETERB1 | First Continuation of the Anglo-Saxon Chronicle E | 1131 |
| 2 | WorcFrag | The First Worcester Fragment | 1135 |
| 3 | TheGrave | The Grave | 1140 |
| 4 | BodySoul | Body and Soul | 1150 |
| 5 | CMPETERB2 | Second Continuation of the Anglo-Saxon Chronicle E | 1154 |
| 6 | CMLAMBET | The Lambeth Homilies* | 1160 |
| 7 | PatNost | Pater Noster | 1160 |
| 8 | CMTRINIT | The Trinity Homilies* | 1160 |
| 9 | CMORM | Ormulum | 1175 |
| 10 | PoemaMorale | Poema Morale | 1175 |
| 11 | ProvAlf | The Proverbs of Alfred | 1175 |
| 12 | CMVICES1 | Vices and Virtues | 1180 |
| 13 | LordOneGod | Lord as Thou art one God | 1195 |
| 14 | Ureisun | A Good Orison of Our Lady | 1200 |
| 15 | WooingGroup | Texts of the Wooing Group | 1200 |

Table 3.3: Overview over the early Middle English text files and their approximate dates of composition

(6) The *Lambeth Homilies* include several sermons that are probably transliterations of $11^{th}$ century, Old English texts. These revisions have not been considered in my studies.

(7) The *Trinity Homilies* contain five sermons that are shared with the *Lambeth Homilies*. Those five copies were removed from the text file.

After all the changes to the corpus had been implemented, there remained a total of 101 electronic text files. They contain approximately 1.5 million words of running text in 96,634 sentence tokens. Figure 3.1 below illustrates their temporal distribution and relative size. As the graphic shows, the chronological order assumed here results in two dense clusters of texts - one around the year 900, associated mainly with translations initiated by King Alfred, such as the Old English *Bede*, *Cura Pastoralis* or *Orosius* etc., the other around the year 1000, related predominately to the writings of Ælfric and Wulfstan. These two peaks in textual production are often used as a convenient criterion for the differentiation between an early and a late Old English sub-period. Their transition is of course fuzzy, but 950 A.D. may serve as a convenient boundary date for their separation. Very few texts survive from the second half of the 11$^{th}$ century on, and almost none from the time immediately following the Norman Conquest in 1066. Hence, it becomes increasingly more difficult to study the evolution of the English language from that time until the onset of substantial Middle English text productions.

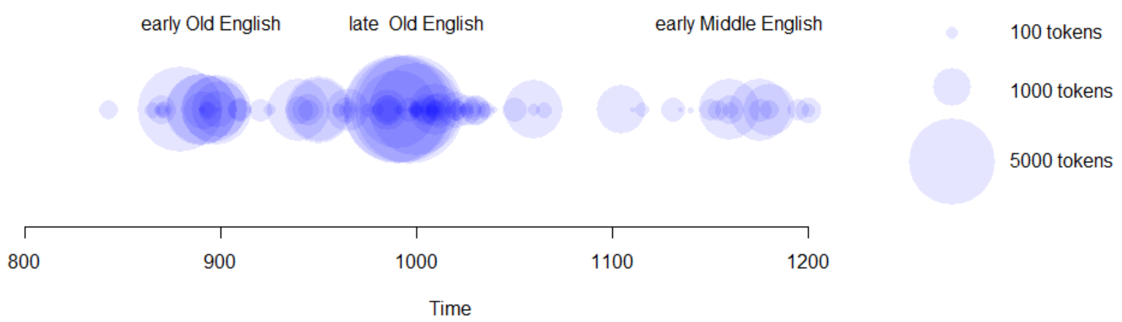



Figure 3.1: Temporal distribution and size of texts considered in this study

All searches described in this chapter were conducted with the software *CorpusSearch* (version 2.002.21) (Randall 2004). For convenience, I will frequently sketch in top captions the search query that an example is based on. In these cases, "X – Y" means that X and Y are immediately adjacent (=CorpusSearch command `iprecedes`), while "X ... Y" means that Y follows X with any number of intervening constituents (possibly zero) (=CorpusSearch command `precedes`).

I will now turn to the construction of a theoretical grammar model of Old English clause structure with special reference to the syntactic status of conjunctions. The corpus files described here will subsequently be employed in the empirical part of this chapter to collect data for the validation of the grammar model.

## 3.3 Modeling Interactions among Old English Syntactic Phenomena

In this first theoretical part of the chapter, I will model interactions between verb placement patterns and topicalization on the one hand and the MC-CC clause type distinction on the other hand. First, I shall develop a grammar model of Old English clause structure (section 3.3.1). Subsequently, I will derive from it specific, quantitatively testable predictions. They concern both synchronic interactions in the form of distributional differences between MCs and CCs as well as diachronic interactions in the form of divergent word order developments in those two clause types (section 3.3.2).

### 3.3.1 A Model of Old English Clause Structure and Syntactic Changes

In this section, I will construct an LFG model of Old English clause structure. It can handle core aspects of the syntax of Old English synchronically and also describes diachronic developments by incorporating competing and hence changing rules. The relevant changes concern the placement of finite verbs and the rate of topicalization. Furthermore, my model will advocate the possibility that Old English conjunctions, too, are in a process of transition at that time, changing from one syntactic category to another.

#### 3.3.1.1 The Syntactic Category of Finite Verbs

Old English finite verbs essentially all pattern alike. In particular, syntactic contexts do not impose differential requirements on finite main verbs as opposed to finite verbs that will eventually develop into modern auxiliaries. Rather, both kinds behave like Present-Day English auxiliaries. I will call the latter class of verbs 'pre-auxiliaries,' or specifically 'pre-modals' for the ancestors of modal auxiliaries, or just 'auxiliaries' for brevity's sake. There are three syntactic contexts that show syntactic identity between Old English finite main verbs and pre-auxiliaries: inversion, VP-adjunct placement and verb phrase ellipsis.

First, both main verbs and pre-auxiliaries can directly invert with the subject, for instance for question formation, as in (3). In the following examples, partial labeled bracketing sketches the parses I would like to assign to the sentences.

(3) a. <u>$V_{main}$ - *Subject*</u>
hwæt [$_C$ [$_I$ **sægest**]] [$_{IP}$ þu us nu be swylcere dæde?]
what say you us now about such deed
'What are you telling us now about such a deed?' (coaelhom,ÆHom_14:189.2110)

b. <u>$V_{preaux}$ - *Subject*</u>
Hwæt [$_C$ [$_I$ **wille**]] [$_{IP}$ we eow swiðor secgan be ðisum symbeldæge.]
what will we you rather say about this feast-day
'What else could we tell you about this holiday?' (cocathom2,ÆCHom_II,_34:259.115.5788)

Secondly, both kinds of finite verbs ordinarily occur before rather than after elements roughly indicating the VP-boundary, such as low adverbs, secondary negation or floating quantifiers. Hence, Old English finite main verbs do not show the word order 'VP-adjunct - finite verb' typical of Modern English main verbs, (4a), but rather the word order pattern 'finite verb - VP-adjunct' as Old English pre-auxiliaries or Modern English auxiliaries, (4b). However, the availability of verb-final structure, postposition and high adjunction of certain light adverbs before the finite verb substantially confound the analysis.

(4) a. <u>*Subject* - $V_{main}$ - *Adv ... Pronoun*</u>
Ic [$_I$ **andette**] [$_{VP}$ eac [$_{VP}$ hine ætforan minum fæder.]]
I acknowledge also him before my father
*'I acknowledge also him before my father'
'I also acknowledge him before my father'
(cocathom2,ÆCHom_II,_43:324.198.7300)

b. <u>*Subject* - $V_{preaux}$ - *Adv ... Pronoun*</u>
we [$_I$ **willað**] [$_{VP}$ eac [$_{VP}$ þæt andgit eow geopenian.]]
we will also that understanding you open
'We will also open up the (proper) understanding (of this passage) for you'
(cocathom2,ÆCHom_II,_12.2:122.414.2672)

Third, there are cases in which finite main verbs can license verb phrase ellipsis in Old English, while this is no longer a possibility today, as shown by (5a). Again, this characteristic is shared with Old English pre-auxiliaries or Modern English auxiliaries, (5b).

(5) a. $\underline{V_{main} \; ... \; V_{main} \; \Delta}$

Drihten [$_{\text{I}}$ **hat**] me gan to ðe upon ðam wætere [...] gif þu [$_{\text{I}}$ **hætst**]. ðonne mæg ic;
Lord order me go to thee upon the water [...] if you order, then may I

'Lord, order me to come to you on the water. [...] If you {*order | do}, then I will be able to.'
(cocathom2,ÆCHom_II,_28:227.189.5048, 5050)

b. $\underline{V_{preaux} \; ... \; V_{preaux} \; \Delta}$

axiað hire gif hi [$_{\text{I}}$ **seo**] frig. and gif hi [$_{\text{I}}$ **is**], þonne wille ic hi habban me to wife;
ask her if she be free and if she is then will I her have me to wife

'Ask her if she is free. And if she is, then I will have her as my wife.
(comargaC,LS_14_[MargaretCCCC_303]:5.6.50, 51)

The inference from these facts must be that Old English finite verbs uniformly belong to the same syntactic category, and, specifically, that their syntactic category corresponds to that of Present-Day English auxiliaries. I shall use the label `I` for 'inflection' for this word class in my model, as in the previous chapter. In derivational models, this formalization could be translated into the categorical endowment of all finite verbs with a feature driving verb movement from `V` to `I`.

Example lexicon entries for finite verbs are shown in (6), where the letter after the graphemic form represents the syntactic category. For illustrative purposes, I also include the verb's subcategorization requirements, a `VFORM` feature that keeps track of their finiteness and grammatical features for tense, mood, subject-verb agreement and case requirements for its arguments.

(6) a. *sægst* I (↑ PRED)= 'SAY <(↑ SUBJ) (↑ OBJ) (↑ OBJ$_{th}$)>'
(↑ VFORM)= fin
(↑ TENSE)= present
(↑ MOOD)= indicative
(↑ SUBJ CASE)=$_c$ nom
(↑ OBJ$_{th}$ CASE)=$_c$ dat
(↑ OBJ CASE)=$_c$ acc
(↑ SUBJ PERS)=$_c$ 2
(↑ SUBJ NUM)=$_c$ sg

b. *wille* I (↑ PRED)= 'WANT/WILL <(↑ SUBJ) (↑ XCOMP)>'
(↑ VFORM)= fin
(↑ SUBJ)= (↑ XCOMP SUBJ)
(↑ XCOMP VFORM)= inf
(↑ TENSE)= present
(↑ SUBJ CASE)=$_c$ nom
(↑ SUBJ NUM)=$_c$ sg
{ (↑ SUBJ PERS)$\neq$ 2
(↑ MOOD)= indicative
| (↑ MOOD)= subjunctive }

#### 3.3.1.2 The Clausal Hierarchy of Projections

Linguists have proposed a great number of models for the description of Old English core syntax. In fact, the puzzles posed by Old English grammar have fascinated formally oriented researchers for a long time (witness, for instance, work such as Johnson 1975). More recently, however, Old English grammar models have become so diverse that it has become increasingly hard to keep track of all the innovations (for selective summaries of recent proposals, see e.g., Burnett and Ferch 2005, Bruening 2016). Further, many different formalisms seem to me to be little more than alternative formulations of a very similar underlying theory. On the one hand, representations suggested by different models often lead to identical constituencies in terms of label-free, bare phrase structure. On the other hand, it seems to be almost impossible to point to straightforward predictions that could differentiate certain proposals from one another. I will briefly present a few examples. For reasons of space, I cannot provide a comprehensive review of motivations, strengths and weaknesses for every proposal. I merely wish to illustrate the recent proliferation of Old English clause structure models and to justify my choice for one of them.

First, a number of scholars work with common projection labels that emerged from the *Government and Binding* framework (e.g., Chomsky 1986), namely CP, IP and VP (core publications: Kroch and Taylor 1997, Pintzuk 1999, Bergen 2003). A parse of a sentence within this hierarchy of projections is shown below.

(7) [$_{CP}$ ælc yfel [$_{C'}$ [$_{IP}$ he [$_{I'}$ mæg [$_{VP}$ don ]]]]]
each evil he may do

'He can do each evil' (WHom, 4.62, parse as explained in Kroch and Taylor 1997, ex. (4a))

This is the approach that I decided to use for my model for the following reasons. (i) It is relatively simple and comprehensible for many linguists. The labels CP, IP and VP are the closest that Old English syntacticians have come to providing a standard terminology. (ii) Nominal subjects are assumed to be variably placed in the specifier of IP or VP following certain information-structural tendencies, the latter option in the process of declining during the Old English period. I find this upward trend for subjects appealing. This analysis is also in accordance with ordinary processes of language change. (iii) The model allows a relatively direct insertion of elements from corpus surface sentences into abstract syntactic trees with the smallest (but still substantial) amount of ambiguity possible. On the downside, it leaves a relatively large amount of unexplained residual data. It seems preferable to me to have some residue that cannot be parsed by a grammar model than to have large amounts of data that can be parsed in very many different ways. (iv) Most LFG-style formalizations of Old English have employed this terminology (e.g., Clark 2004, Mahowald 2011).

Secondly, many models postulate not one but two functional projections between CP and VP (core publications: AGR1P and AGR2P in Cardinaletti and Roberts 1991, Haeberli and Haegeman 1995; AgrSP and TP in Haeberli 2000; FP and TP in Kemenade 1999 and most of her subsequent work as well as Fischer et al. 2000; TP and MP in Speyer 2010). An example analysis is shown in (8).

(8) [$_{CP}$ hiora untrymnesse [$_{C'}$ [$_{AgrSP}$ he [$_{AgrS'}$ sceal [$_{TP}$ [$_{VP}$ ðrowian on his heortan ]]]]]]
their weakness he shall atone in his heart

'He shall atone for their weakness in his heart' (CP,60.17, parse as explained in Haeberli 2000, ex. (2))

The following reasons led me to discard this approach for my formalization. (i) For the vast majority of cases it seems to be merely a notational variant of the previous option. For instance, the constituent structures in (7) and (8) are almost identical; differences arise only through different labels and an additional vacuous projection. Arguments adduced for the necessity for a second functional projection above VP are extremely intricate and maybe inconclusive. (ii) The higher functional projection is usually assumed to become inactive during Middle English. The resulting downward shift in subjects and subsequent ambiguity in verb placement does not seem convincing to me. It does not constitute a well-understood process of language change. (iii) The additional structure provided by an additional functional projection can be exploited to make better sense of frequency asymmetries between matrix and embedded clauses in verb placement (Haeberli 2005) and the distribution of certain adjuncts such as the (discourse particle) adverbs *þa, þonne* 'then' (Kemenade 2009), the (secondary or emphatic) negative adverb *na* (Kemenade 2011) or adverbs in general (Haeberli and Ingham 2007). However, the potentially wider coverage of facts comes at the, in my opinion unacceptably high, cost of increased structural ambiguity. This makes these models difficult to use as a basis for the descriptions of Old English and therefore very hard to falsify.

Thirdly, a number of scholars have adopted a Split-CP model (Rizzi 1997) for Old English (core publications: Hinterhölzl and Petrova 2010, Walkden 2014, Gelderen 2017 for Split-CP in main clauses only). An illustrative parse is presented bellow.

(9) [$_{ShiftP}$ hiora untrymnesse [$_{Shift'}$ [$_{FamP}$ he [$_{Fam'}$ [$_{FinP}$ sceal [$_{TP}$ ðrowian ]]]]]]
their weakness he shall atone

'He shall atone for their weakness' (CP,60.17, parse as explained in Walkden 2014, ch. 3, ex. (41))

I decided against splitting the CP in my model because of the following considerations. (i) The resulting constituencies are once again largely identical to the previous proposals. Examples (7), (8) and (9) all have the same dominance relations; merely the labels and the number of vacuous projections are different. (ii) Split functional projections allow for convenient descriptions of different kinds of discourse functions (e.g., aboutness topics, familiar topics, contrastive focus etc.) and enough structure for embedded topicalization. However, these potential advantages must be weighed against increased structural ambiguity, a certain clumsiness in the posited changes from Old to Modern English and the fact that these approaches have not yet explored central aspects of Old English grammar, such as the distribution of adverbial sentence elements or weak pronouns. (iii) I find the across the board equation of the discursive role of a sentence element with syntactic positions or syntactic features conceptually ill-conceived. The information structural status of an element is usually an emergent phenomenon, i.e., a consequence of its introduction into a discourse, and only exceptionally a theoretical primitive of syntax, i.e., a position such as `TopP` or a formal feature like `[+Topic]`.

Fourth, there are formalizations of Old English clause structure incorporating Kayne's Linear Correspondence Axiom (Kayne 1994) and a large number of movement operations (core publications: Biberauer and Roberts 2005 and subsequent work, Wallenberg 2009). Example (10) shows a derivation for a sentence with the order 'main verb - auxiliary' which involves pied-piping of an entire *v*P to the specifier of `TP`, indicated by the index $_k$.

(10) þa [$_{\text{TP}}$ [$_{v\text{P}}$ se Wisdom [$_{v\text{P}}$ [$_{\text{VP}}$ þa t$_i$ þis fitte]$_j$ asungen$_i$+$_v$ t$_j$ ]]$_k$ [$_{\text{T}}$ hæfde] t$_k$ ]
when the wisdom then this poem sung had
'When Wisdom had sung this poem' (Boeth 30.68.6, parse as in Biberauer and Roberts 2005, ex. (12))

I am somewhat more skeptical of such approaches than the others and so did not consider using them as a basis for my formalization. I will direct my criticism at Biberauer and Roberts (2005) specifically, but many points will carry over to similar proposals. (i) There are numerous practical issues. The model leaves a large number of important mechanisms of Old English unexplored, such as case assignment, adverb placement, possible determinants of the optional movement operations etc. The model also results in great structural ambiguity, especially once additional operations for the generation of the left periphery are taken into account. As a consequence, it cannot easily be used as a basis for the description of and empirical research on Old English word order patterns. (ii) More abstractly speaking, the model rests on a fairly large number of inherently unfalsifiable assumptions. Only the representations generated by derivational processes can be probed, not the kinds of posited movement operations themselves. (iii) To the extent that the Kaynean model generates representations that are different than the ones from the previous formalizations, the resulting constituency relations often appear to be incorrect. I shall give just one example.

Biberauer and Roberts' system entails that subject and main verb in (10) form a constituent, `vP`, to the exclusion of the clause-final finite verb, [[$_{v\text{P}}$ SV]Aux]. Most traditional models, on the other hand, would predict that the verb phrase and the auxiliary form a constituent, say `I'`, to the exclusion of the subject, [S[$_{\text{I}'}$VAux]]. While it is indeed difficult to decide between these two options, I found three constructions that seem to necessitate the existence of a constituent containing the verb phrase and the finite verb.

First, the verb phrase material and the auxiliary can be coordinated with across the board extraction of the subject, as in (11). Such constructions are unexpected under Biberauer and Robert's model since the conjoined material is not assumed to form a constituent and hence cannot be targeted by coordination. In contrast, traditional approaches can easily parse such sentence by assuming that the coordinated phrase corresponds to the node dominating the verb phrase and the auxiliariy, i.e., `I'`.

(11) ... ðonne we nohwæðer ne [$_{\text{I}'}$ hit witan nyllað] ne [$_{\text{I}'}$ hit betan nyllað]
... when we neither it know not-want nor it improve not-want
'... when we neither want to know it nor want to improve it'
(cocura,CP:28.195.5.1302)

Next, stripping is a possible form of ellipsis in Old English, as shown in (12). Biberauer and Roberts are likely to find such sentences difficult to generate because the elided material straddles a string that does not correspond to a constituent in their model. They might thus introduce yet another operation that vacates the subject before *v*P-pied piping. Alternative proposals can analyze such structures without further ad-hoc repairs by truncating the constituent that corresponds to the clause excluding the subject, i.e., `I'`.

(12) Gif [[$_{\text{DP}}$ hors] [$_{\text{I}'}$ gescoten sy]] oððe [[$_{\text{DP}}$ oþer neat] [$_{\text{I}'}$ $\Delta$ ]]
if hors shot be or other animal
'If a horse is shot, or another animal (is shot)'
(colacnu,Med_3_[Grattan-Singer]:118.1.575)

Finally, Biberauer and Roberts might run into problems with 'main verb - pre-auxiliary' sentences that also involve some form of subject raising. A straightforward example of object-to-subject raising in a modal passive is shown in (13).

(13) HU [$_{\texttt{DP}}$ DÆGREDSANGAS] [$_{\texttt{I}'}$ ON FREOLSTIDUM $\texttt{PRO}_{\texttt{arb}}$ _ TO HEALDENNE SYN].

how morning-songs in festival-time to hold are
'How one should perform morning songs during the time of a festival'
(cobenrul,BenR:12.36.9.494)

The element *dægredsangas* 'morning songs' functions as the object of *to healdenne* 'to hold' and as the subject of *syn* 'are,' as informally sketched by an arrow. Its subject status becomes evident from the fact that it triggers subject-verb agreement in plural number. Presumably, Biberauer and Roberts would let the subject raise to the specifier of `TP` (unfortunately, subject-verb agreement is one of the mechanisms not described in Biberauer and Roberts 2005 so that the exact generation of the sentence remains vague). Since this position is now occupied by a subject `DP`, it cannot also be filled by the whole pied-piped *vP* at the same time. The order 'verb - auxiliary' then becomes impossible to derive. Non-Kaynean models, on the other hand, can simply assume that the subject is placed in the high subject position and that the `VP` and finite verb form a constituent to the exclusion of the subject, i.e., `I'`.

Having justified my choice for a particular approach to the description of Old English syntax, I shall now continue to implement an LFG toy grammar for my purposes. The backbone of the Old English clause can be analyzed with phrase structure rules similar to those for Modern English, as laid out, for instance, in the previous chapter. This allows for continuity in the structure of the clausal spine throughout the history of English. I will focus on finite main clauses and use the label $\texttt{CP}_{MAT}$ as the start symbol for all matrix clauses.

The specifier of a matrix clause `CP` can remain unprojected so that the start symbol simply re-writes to `C'`, thereby imposing the constraint that the clause be finite, (14). Fronting of a function to the specifier of `CP` will be discussed separately in the section on topicalization under 3.3.1.6.

(14) $\texttt{CP}_{MAT} \rightarrow$ $\texttt{C'}_{MAT}$
↑ = ↓
(↓ TENSE)

Outside of fairly well-defined inversion contexts in Old English, $\texttt{CP}_{MAT}$ does not project a verbal head. Instead, `C'` will directly re-write to the complement `IP`. By doing so, the clause will be typed as declarative, `decl`, (15). The contexts in which the verb appears under `C` will be discussed more fully in the section on high verb placement under 3.3.1.4.

(15) $\texttt{C'}_{MAT} \rightarrow$ IP
↑ = ↓
(↑ CLAUSE-TYPE) = decl

The specifier of `IP` can be a subject position, as in Modern English, (16a). However, it can also remain unprojected, (16b). Alternatively, one could categorically introduce a null subject expletive, as has been proposed numerous times, often in a way that simply takes such an analysis for granted (e.g., Kemenade 1997, Fischer et al. 2000: 71, Williams 2000, Haeberli 2002*a*, Gelderen 2013, Ringe and Taylor 2014: 502). I see little reason to introduce empty elements for purely theoretical reasons and would accept them only if they were absolutely needed to account for empirical data. This does not seem to be the case here.

(16) a. IP → DP I'
(↑ SUBJ) = ↓ ↑ = ↓
b. IP → I'
↑ = ↓

The `I'`-level will project a verbal head, `I`, in ordinary declarative clauses. `I` may select `VP` to the right as its complement, (17a). Alternatively, the hierarchy of projections may terminate at this point, for instance for intransitive predicates, (17b). For now, the rule below can only handle finite verbs placed in front of the `VP`. `I`-final headedness will be implemented in the section on variation between verb-medial and verb-final structures under 3.3.1.3.

(17) a. I' → I VP
↑ = ↓ ↑ = ↓
(↑ CLAUSE-TYPE) $=_c$ decl
b. I' → I
↑ = ↓
(↑ CLAUSE-TYPE) $=_c$ decl

The specifier of VP is modeled as a low subject position. It is not a new insight that Old English requires such an additional position for subjects. First of all, the element order 'verb-subject' is found, albeit rarely, in subordinate clauses and thus requires a place for the subject after the complementizer C and the finite verb I (e.g., Mitchell 1985: §3926, and related §3101). A few examples are shown in (18).

(18) *Complementizer - V - Subject*

a. [$_{CP}$ Gif [$_{IP}$ sie [$_{VP}$ [$_{DP}$ þære adle bryne] innan þæs strang þæt mon ne mæge utan geseon]]]
if be the disease's burning inside that strong that one not may outside see
'If the burning of the diseases is strong to such a degree that one cannot see it from the outside'
(colaece,Lch_II_1:4.4.10.487)

b. [$_{CP}$ swa [$_{IP}$ wæs [$_{VP}$ [$_{DP}$ woruldbebod] gebannen fram Aguste þam casere]]] swa wæs eft ...
as was world-command decreed from Augustus the Caesar so was also ...
'Just as the universal edict was decreed by emperor Augustus, so was also ...'
(coverhom,HomS_1_[ScraggVerc_5]:45.874)

c. ... [$_{CP}$ buton þæt [$_{IP}$ secgað [$_{VP}$ [$_{DP}$ þa ealdan gesetnissa] heora yfelan dæda]]]
... except that say the old laws their evil deeds
'... except for the fact that the Old Testament tells of their evil deeds'
(colsigewZ,ÆLet_4_[SigeweardZ]:8.8)

More importantly, however, the nature of verb-second in Old English also suggests a lower subject position. It has been known for a long time that verb-second patterns are systematically disrupted by subject pronouns intervening between relatively heavy, clause-initial topics, say DPs or PPs, and the finite verb. The resultant order 'topic - subject pronoun - verb' is typical of Old English, (19a). In contrast, non-pronominal (also referred to as full or nominal) subjects in such topic-initial clauses can occur in front of the verb, (19b), but also after it, (19c). Pronominal subjects never occur after the verb in this context, (19d), (e.g., Roth 1914, Fourquet 1938, Canale 1978; for an overview over studies on Old English constituent order from its beginnings to the 1990s, see Davis 2006: 74-83). Yet, it was not before the 1990s that a coherent analysis of these facts was developed (most closely related to my account in Kroch and Taylor 1997, building on material published in Pintzuk 1999 and Pintzuk 1993). Essentially, the word order patterns can be modeled elegantly and economically by assuming that subject pronouns must necessarily be positioned in the high subject position, the specifier of IP, whereas non-pronominal subjects can occur in the high subject position, the specifier of IP, but also in a lower subject position, the specifier of VP. In one way or another, the majority of scholars have adopted such an analysis incorporating a secondary lower subject position to account for the Old English verb-second constraint.

(19) a. *Topic - Pronominal subject - V*

[$_{CP}$ þis bebod [$_{IP}$ [$_{DP}$ ic] *nam* [$_{VP}$ æt minum Fæder ]]]
this command I took at my father
'I received this command from my father'
(cowsgosp,Jn_[WSCp]:10.18.6634)

b. *Topic - Non-pronominal subject - V*

[$_{CP}$ þas cyþnesse [$_{IP}$ [$_{DP}$ Drihten] *nam* [$_{VP}$ of þisse wisan ]]]
those testimonies Lord took of this manner
'The Lord took those (scriptural) testimonies in this manner'
(coblick,HomS_10_[BlHom_3]:31.81.416)

c. *Topic - V - Non-pronominal subject*

[$_{CP}$ ðæne aþ [$_{IP}$ *nam* [$_{VP}$ [$_{DP}$ Wulfsige se scirigman] ]]]
the oath took Wulfsige the sheriff
'Sheriff Wulfsige received the oath'
(codocu3,Ch_1458_[Rob_41]:36.50)

d. * *Topic - V - Pronominal subject*

There are two plausible ways to implement the ban on the placement of pronominal subjects in the low subject position, the specifier of VP. The first option would be to annotate this position with a negative equation that directly rules out the presence of an element that has the formal property of being a pronoun.

This approach seems somewhat crude, but formalizes the actual observation. The second option would be to model the lower subject position as a an information-structurally constrained position. For instance, if the position maps onto a discourse-new or focus function, the required information status would be incompatible with the anaphoricity of pronouns. In this case, pronominal subjects would not technically be ungrammatical in the lower subject position but merely receive an infelicitous focus interpretation there. The tendency that low subjects are discourse-new, generic or non-referential does in fact seem valid for Old English, at least when measured with the adverbs (or discourse particles) *þa*, *þonne* 'then' as an indication of the boundary between the higher and the lower subject position (e.g., Kemenade 2009). This approach seems more elegant, but faces two problems. First, it is very difficult to delimit the information-structural status of a sentence element objectively. Secondly, there are counterexamples to the tendency that low subjects must be discourse-new. Problematic cases are shown for main clauses in (20a) and for subordinate clauses in (20b) below.

(20) a. And þa ofer þone midne sumor com **þa** [$_{\text{VP}}$ [$_{\text{DP}}$ se micla flota] to Sandwic]
and then over the mid summer came then the great fleet to Sandwich
'Then, over midsummer, the great fleet came to Sandwich'
(ChronC_[Rositzke]:1006.4.1387)
(the great, i.e., Danish, fleet is mentioned repeatedly in the annals immediately preceding this one, and so must be regarded as discourse-old)

b. þa þæt **þa** [$_{\text{VP}}$ [$_{\text{DP}}$ se his gefera] geseah & ongeat], þa frægn he hine
when that then the his companion saw and perceived, then asked he him
'When his companion saw that, then he asked him'
(cobede,Bede_4:26.352.22.3553)
(discussed as a counterexample to generalization that low subjects must be discourse-new in Kemenade et al. 2008: 14, ex.(25))

These problems and the fact that I would like to keep my toy grammar close to the observations it is based on led me to implement a version of the first option.

How then can the specifier of VP be made sensitive to pronominal elements? The value of the PRED feature of pronouns can be represented as PRO. An example lexicon entry for the pronoun *ic* 'I' is shown in (21).

(21) *ic* D (↑ PRED)= 'PRO'
(↑ PRON-TYPE)= personal
(↑ CASE)= nom
(↑ PERS)= 1
(↑ NUM)= sg

The functional annotations on the specifier of VP can then simply preclude a function whose PRED feature is valued as PRO. The resultant phrase structure rules developing the VP-node are shown in (22). The rule in (22a) re-writes VP to a non-pronominal subject in the specifier of VP and V′ thus capturing cases in which a low subject is followed by verb phrase material (e.g., (18), (20a)). The rule in (22b) re-writes VP to a non-pronominal subject in the specifier of VP only and can thus parse structures in which the low subject may terminate the clausal spine (e.g., (19c)). Finally, the rule in (22c) continues VP with V′ for clauses in which the specifier of VP is empty (e.g., (19a), (19b)).

(22) a. VP → DP V′
(↑ SUBJ) = ↓ ↑ = ↓
(↓ PRED) ≠ PRO

b. VP → DP
(↑ SUBJ) = ↓
(↓ PRED) ≠ PRO

c. VP → V′
↑ = ↓

The V′-node can be re-written in a fair number of different ways to account for all the different kinds of Early English verbal complements. For illustrative purposes, I shall present just two possibilities - the verb phrase might contain an object DP for transitive verbs (23a) or a non-finite clause to host the XCOMP-complement of pre-modals (23b), etc.

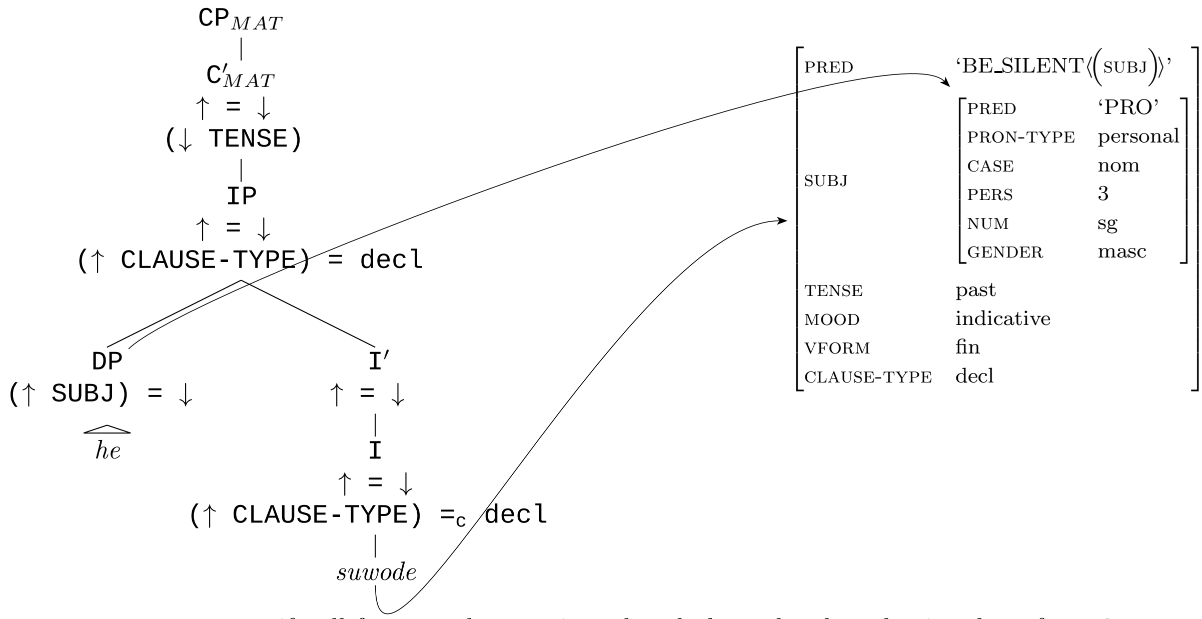
(23) a. V′ → DP
(↑ OBJ) = ↓

b. V′ → IP
(↑ XCOMP) = ↓

etc.

This concludes my description of the fundamental hierarchy of projections in Old English. I will end this section with an illustration of the workings of my grammar model at this stage with a simple 'pronominal subject - verb' structure, such as *He suwode* 'He remained silent.'

(24) He suwode
he was-silent
'He remained silent' (cowsgosp,Mk_[WSCp]:14.61.3399)




The parser attempts to unify all features that are introduced along the clausal spine, here from $\texttt{CP}_{MAT}$ to I. The matrix CP-layer requires that the clause be tensed. The finite main verb *suwode* values the tense attribute as past and hence the finiteness requirement is met. Since C′ re-writes directly to IP, the finite clause must be declarative. The clause type is also checked by the verb in I. The main verb subcategorizes for a subject, which is lexically constrained to be nominative and either first / third person singular in the indicative mood or any person singular in the subjunctive mood. Here, the pronoun *he* is interpreted as a subject by virtue of being placed in the specifier of IP and meets all of the required morpho-syntactic features. I do not analyze the internal constituent structure of DPs but use triangles instead. The model permits one solution for the parse of this word order (strictly speaking two, one for indicative, the other for subjunctive mood of the lexically ambiguous verb).

#### 3.3.1.3 Verb-Medial and Verb-Final Structures

The finite verb in Old English is placed under I by default. By that I mean that the verb occurs in I in ordinary, positive, indicative declaratives clauses without a clause-initial operator. Basic rules for such ordinary declarative clauses were implemented in (17). However, in addition to the rule that places I in front of VP as in Modern English, the grammar must also contain a phrase structure option for the generation of VP in front of I as, say, in Modern Dutch. Old English grammar can thus generate IPs that are head-initial (traditionally called verb-medial, *SVX*) as well as head-final (traditionally called verb-final, *SXV*). This is true for main as well as subordinate clauses, but the former are far less likely to show verb-final patterns than the latter. The additional phrase structure rule is presented in (25) below.

(25) I′ → VP I
↑ = ↓ ↑ = ↓
(↑ CLAUSE-TYPE) $=_c$ decl

The variation between head-initial and head-final IPs can be illustrated best with elements that least ambiguously count as VP-internal material. Nominal objects, for instance, have a good chance of indicating the position of VP if they are relatively light. Thus, the order ‘object - verb’ in (26a) is very likely parsed with the I-final phrase structure rule in (25), while the order ‘verb - object’ in (26b) is plausibly generated by the I-initial phrase structure rule in (17a). However, the object in the second example might also have postposed. This option will be implemented in the section on extraposition under 3.3.1.5.

(26) a. *I-final headedness*

... þa [$_{\text{IP}}$ ðu [$_{\text{I}'}$ [$_{\text{VP}}$ **mæstne welan** ] *hæfdest* ]]
... when you most wealth had

’... when you had the greatest wealth’ (coboeth,Bo:26.58.24.1078)

b. *I-initial headedness (or extraposition of the object)*

... for ðan ðe [$_{\text{IP}}$ we [$_{\text{I}'}$ *habbað* [$_{\text{VP}}$ **heofenlice welan** ]]]
... for that that we have heavenly wealth

’... because we have heavenly wealth’ (cocathom2,ÆCHom_II,_38:281.63.6343)

Similarly, object pronouns can be placed within the VP. Hence, the order ‘object pronoun - verb’ is potentially parsed as an I-final structure, (27a). However, it is also possible that the object pronoun preposes to a structurally distinct position between subject and verb. I will consider this option in the section on pronominal preposing under 3.3.1.7. Since pronouns cannot extrapose, the order ‘verb - object pronoun’ must be analyzed as an I-initial clause, (27b).

(27) a. *I-final headedness (or preposing of the pronoun)*

[$_{\text{IP}}$ he [$_{\text{I}'}$ [$_{\text{VP}}$ **hig** ] *þreade* ]]
he them rebuked

‘He rebuked them’ (cowsgosp,Lk_[WSCp]:9.55.4439)

b. *I-initial headedness*

[$_{\text{IP}}$ Hi [$_{\text{I}'}$ *sohton* [$_{\text{VP}}$ **hyne** ]]]
they sought him

‘They sought him’ (cowsgosp,Mt_[WSCp]:21.46.1488)

The most reliable diagnostics for VP are perhaps separable particles such as *ut* ‘out’ if their meaning is semantically transparent (as defined, for example, in Wasow 1997 and Hawkins 2000). These items neither postpose as nominal objects might; nor do they ordinarily prepose out of the VP as pronominal objects frequently do. Therefore, the order ‘particle - verb, ’ as in (28a), very likely indicates a head-final, the order ‘verb - particle,’ as in (28b), a head-initial, IP. The first example is unlikely to have the particle placed under I along with the finite verb because the verb already occurs with a prefix, *a-drifan*, with some sort of telic meaning, ‘drive away, drive out, expel’.

(28) a. *I-final headedness*

manega deofolseocnyssa [$_{\text{IP}}$ he [$_{\text{I}'}$ [$_{\text{VP}}$ **ut** ] *adraf* ]]
many devil-sicknesses he out drove

‘He drove out many demoniacal possessions’ (cowsgosp,Mk_[WSCp]:1.34.2248)

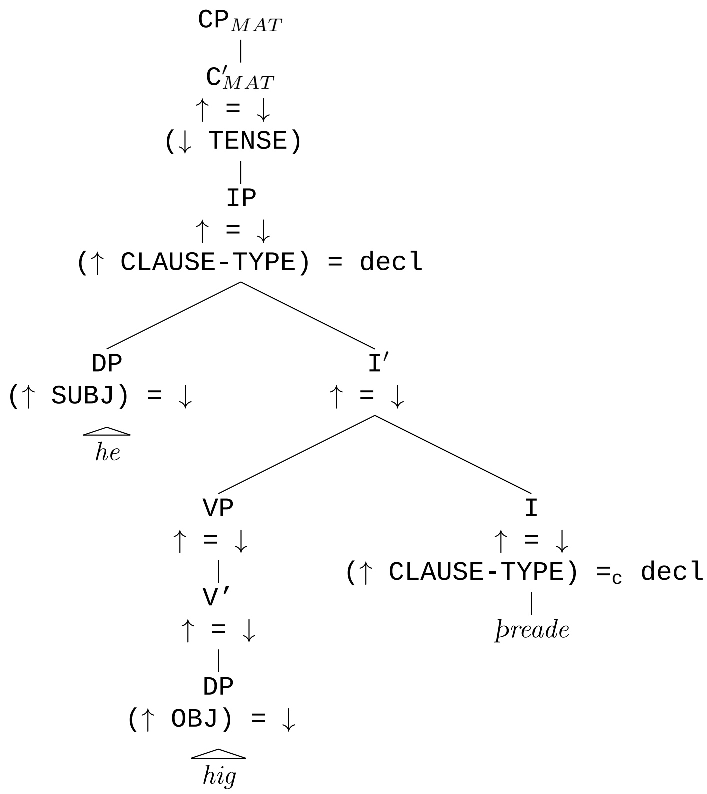
b. *I-initial headedness*

þa yflan [$_{\text{IP}}$ hig [$_{\text{I}'}$ *awurpon* [$_{\text{VP}}$ **ut** ]]]
the evil they threw out

‘They threw out the bad ones’ (cowsgosp,Mt_[WSCp]:13.48.900)

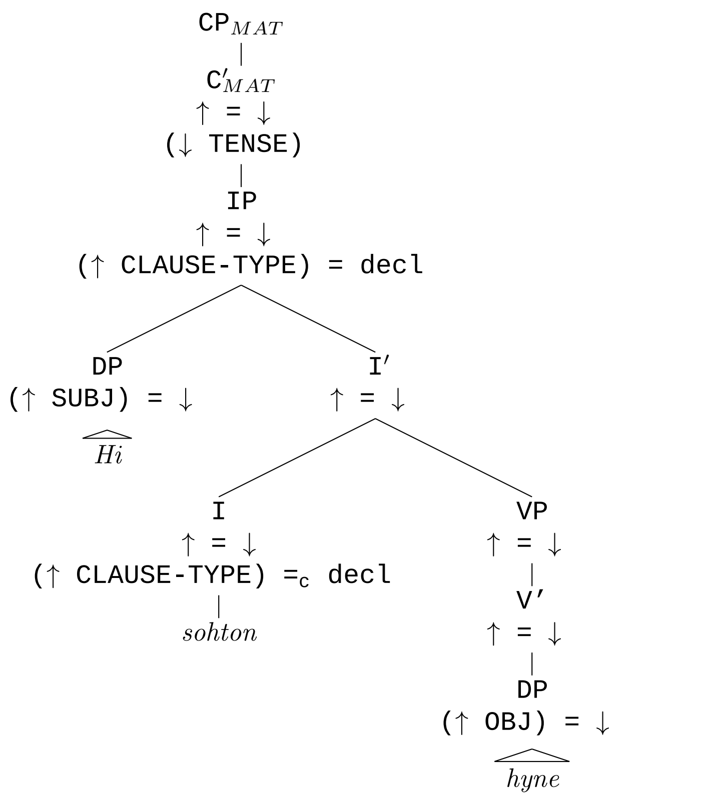
To make my analysis maximally explicit, I will present the full parses for the two examples in (27) as an illustration in the trees below.

(29) he hig þreade
he them rebuked
'He rebuked them' (cowsgosp,Lk_[WSCp]:9.55.4439)




| | | |
|---|---|---|
| PRED | 'REBUKE⟨(SUBJ)(OBJ)⟩' | |
| SUBJ | PRED | 'PRO' |
| | PRON-TYPE | personal |
| | CASE | nom |
| | PERS | 3 |
| | NUM | sg |
| | GENDER | masc |
| OBJ | PRED | 'PRO' |
| | PRON-TYPE | personal |
| | CASE | acc |
| | PERS | 3 |
| | NUM | sg |
| TENSE | past | |
| MOOD | indicative | |
| VFORM | fin | |
| CLAUSE-TYPE | decl | |

(30) Hi sohton hyne
they sought him
'They sought him' (cowsgosp,Mt_[WSCp]:21.46.1488)




| | | |
|---|---|---|
| PRED | 'SEEK⟨(SUBJ)(OBJ)⟩' | |
| SUBJ | PRED | 'PRO' |
| | PRON-TYPE | personal |
| | CASE | nom |
| | PERS | 3 |
| | NUM | pl |
| OBJ | PRED | 'PRO' |
| | PRON-TYPE | personal |
| | CASE | acc |
| | PERS | 3 |
| | NUM | sg |
| | GEND | masc |
| TENSE | past | |
| MOOD | indicative | |
| VFORM | fin | |
| CLAUSE-TYPE | decl | |

The declarative matrix `CP`s are functionally headed by finite verbs, which subcategorize for a subject and an object argument and impose constraints on their case realization as well as subject-verb agreement in person and number. The subject pronoun must be placed in the specifier of `IP` and the pronominal object can only be parsed in the complement position of `VP` within the assumptions of my formal grammar model at the present stage. The pre-verbal object pronoun must thus be parsed with the phrase structure rule creating `I`-final structures, (25), the post-verbal pronominal object with the one for `I`-initial structures, (17a). The model arrives at one unique solution for each word order (although technically more because of syncretism in the mood of *þreade* and the fact that *hig* can be a third person plural pronoun 'them' or a third person singular feminine pronoun 'her').

The variation between `I`-final and `I`-initial structures is indicative of a change in progress. The conservative verb-final positioning is gradually replaced by innovative verb placement as in Modern English. I interpret this to mean that the rule for the generation of `[VP I]` structures competes with and is eventually ousted by the rule for the generation of `[I VP]` orders. It is noteworthy that Old English shows variation between the two verb placement options from its earliest text witnesses on, i.e., the pure, conservative state is not attested.

It seems fair to say that the above development constitutes the best researched change in the literature on Early English syntax. Several scholars could successfully demonstrate an increase in word order patterns reminiscent of Modern English with a constituent after subject and verb, *SVX*, relative to structures only found in medieval English with a constituent between subject and verb, *SXV* (notable publications: Smith 1893, Saitz 1955, Kohonen 1978). The 1990s saw a sharpening of these measurements. Not just any constituent, *X*, was considered as an acceptable `VP`-element because *SVX* sentences might be parsed as extraposition structures. Instead, scholars focused on items that would necessarily identify the position of `VP` in post-verbal position. Hence, the proportion of 'subject - verb - diagnostic' vs. 'subject - diagnostic - verb' orders will now reliably measure a lower bound of underlyingly `I`-initial headedness. These diagnostic elements are pronouns as in the illustrations above, (27), particles, as in (28) (Pintzuk 1999), negatively quantified objects (e.g., Pintzuk 2005) and stranded prepositions (e.g., Pintzuk and Haeberli 2008). I would add to this list pronouns that co-occur with reflexive, intensifying or pronominally reinforced *self* (Zimmermann 2012).

The change is far more advanced in main than in subordinate clauses. It is unknown why there is such a substantially different base probability of finding innovative `IP`-headedness between matrix and embedded clauses. Both formal and functional factors may play a role. Base probabilities also differ with respect to different diagnostic elements that are used to measure the change. What seems clear, however, is that the change proceeds at the same rate of change in all contexts, i.e., the change obeys the Constant Rate Hypothesis. For example, Pintzuk (1999: 220-4) uses a VARBRUL analysis to show that verb placement is innovated equally fast in main and subordinate clauses. Her graph is presented in figure 3.2 below. One can clearly discern that the lines for the two clause types appear to move parallel to each other and hence develop at the same speed.

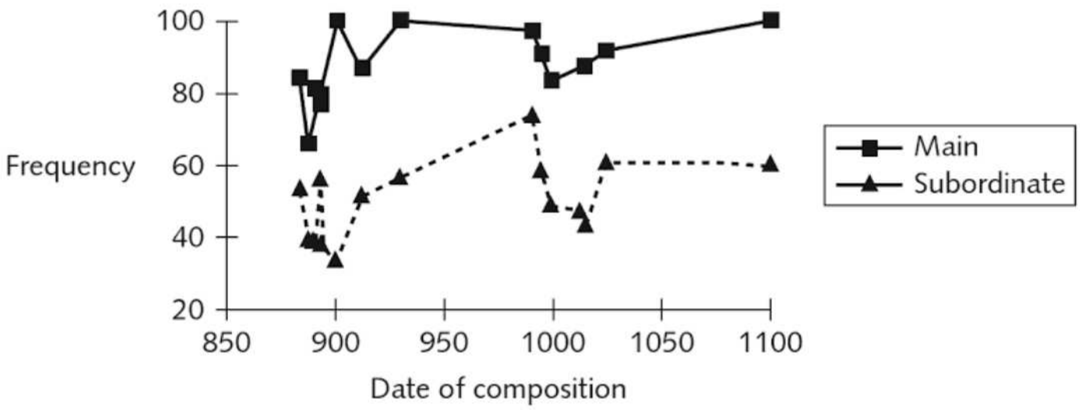
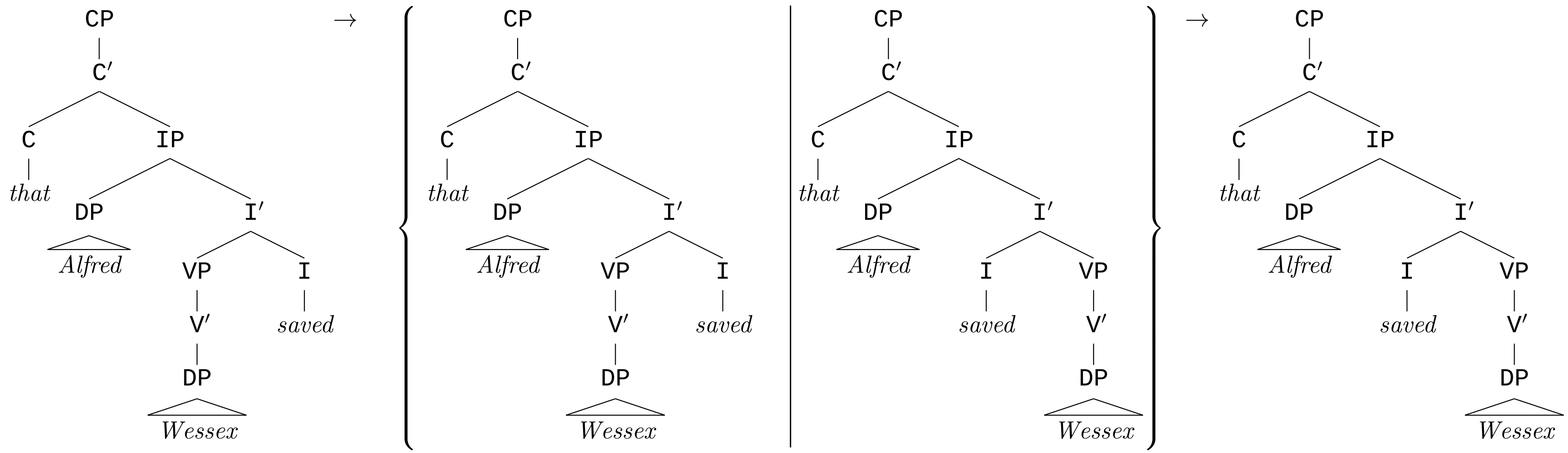



Figure 3.2: Frequency of I-Initial structure in Old English clauses with auxiliary verbs, 884-1100
From: Pintzuk (1999: 220), as reprinted in Janda and Joseph (2003*b*: 523)

The graphic below summarizes the change in the headedness of `IP`. It is formulated in terms of the general model of the implementation of a linguistic innovation in a group of speakers presented in chapter 1.

**Change in Verb Placement: IP-Headedness**

| Conservative period: Speakers represent only the conservative variant $\alpha$ | Transitional period: Speakers represent both the conservative variant $\alpha$ and the innovative variant $\beta$ | Innovative period: Speakers represent only the innovative variant $\beta$ |
|---|---|---|
| $\alpha \rightarrow$ | $\{\ \alpha \mid \beta\ \} \rightarrow$ | $\beta$ |
| Conservative period: Speakers represent only I-final headedness (pure conservative state not attested in the history of English) | Transitional period: Speakers represent both I-final and I-initial headedness | Innovative period: Speakers represent only I-initial headedness |



#### 3.3.1.4 Contexts Triggering High Verb Placement

Old English finite verbs can be positioned in the higher head position, `C`, in a small number of fairly well-defined contexts. Since subject pronouns unvaryingly occur in a position below `C`, the specifier of `IP`, the verb will consistently precede pronominal subjects in these select environments. In the following sections, I will formalize the four most widely recognized contexts triggering high verb placement.

(1) **Interrogatives**. The first of the contexts inserting verbs under `C` are matrix clauses with interrogative force. They may be either yes/no-questions, (31a), or *wh*-questions, (31b).

(31) a. [$_{CP}$ [$_{C}$ [$_{I}$ *Com* ]] [$_{IP}$ he to heom gescrydd ?]]
came he to them clothed
'Did he come to them with clothes on?' (coeluc2,Eluc_2_[Warn_46]:9.10)

b. [$_{CP}$ hwanon [$_{C}$ [$_{I}$ *come* ]] [$_{IP}$ þu Giezi ?]]
whence come you Gehazi
'Where do you come from, Gehazi?' (cocathom1,ÆCHom_I,_27:408.241.5443)

Verb placement under `C` in matrix questions is categorical in Old English. The order 'subject - verb' cannot normally be used to express a direct question. Furthermore, this form of question formation must be regarded as old, potentially inherited from Proto-Germanic as it can already be encountered in Gothic (Fuss 2003) and will remain stable throughout the history of English up to the present day ("residual V2", Rizzi 1996). It is therefore required that my grammar should make `C` the mandatory position for verbs in this context.

I will focus here on modeling yes/no-questions. They can be formalized as follows. I will assume that matrix `C'` can re-write to the head, `C`, and its complement, `IP`, if at the same time the clause type feature is valued as interrogative, (32). The mood of the verb must be indicative in this case since only indicative verbs can form direct questions.

(32) $\mathtt{C'}_{MAT} \rightarrow \begin{array}{c} \mathtt{C}_{MAT} \\ \uparrow = \downarrow \\ (\downarrow \mathtt{MOOD}) =_c \mathtt{indicative} \end{array} \quad \begin{array}{c} \mathtt{IP} \\ \uparrow = \downarrow \\ (\uparrow \mathtt{CLAUSE\text{-}TYPE}) = \mathtt{interrog} \end{array}$

The head of a finite matrix clause, $\mathtt{C}_{MAT}$, will then be re-written as a finite verb of category `I`, (33). This rule is often referred to as 'I-to-C' or 'V-to-C movement' in derivational frameworks.

(33) $\mathtt{C}_{MAT} \rightarrow \begin{array}{c} \mathtt{I} \\ \uparrow = \downarrow \end{array}$

Where the finite verb appears under `C`, it should not be allowed to occur under `I` at the same time. My model therefore requires an additional rule for the production of `I'` that re-writes the node directly to `VP` provided that the clause-type is not declarative, (34a). Alternatively, the clause structure may end after the introduction of a subject in the specifier of `IP`, (34b).

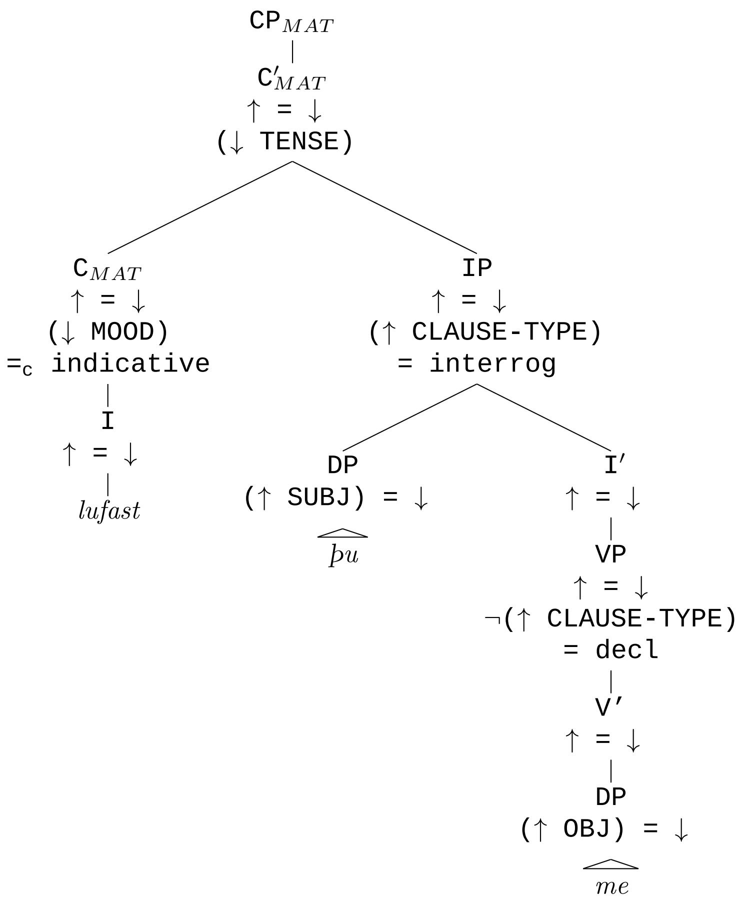
(34) a. $\mathtt{I'} \rightarrow \begin{array}{c} \mathtt{VP} \\ \uparrow = \downarrow \\ \neg(\uparrow \mathtt{CLAUSE\text{-}TYPE}) = \mathtt{decl} \end{array}$

b. $\mathtt{IP} \rightarrow \begin{array}{c} \mathtt{DP} \\ (\uparrow \mathtt{SUBJ}) = \downarrow \end{array}$

A result of this model set-up is that interrogative force can only be introduced if a `C`-head is projected. This captures the obligatory nature of verb placement under `C` in interrogative clauses. Furthermore, these rules are essentially identical to the ones needed for the same context in Modern English, as discussed around example (42) in the previous chapter. Hence, my model also implies some continuity in question formation throughout the history of English.

The Old English toy grammar can now analyze direct matrix questions. The tree in (35) illustrates its workings for the question *lufast þu me* 'Do you love me?'.

(35) lufast þu me
love you me
'Do you love me?' (cowsgosp,Jn_[WSCp]:21.17.7493)




$$\left[\begin{array}{ll}
\text{PRED} & \text{`LOVE}\langle(\text{SUBJ})(\text{OBJ})\rangle\text{'} \\
\text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{`PRO'} \\ \text{PRON-TYPE} & \text{personal} \\ \text{CASE} & \text{nom} \\ \text{PERS} & 2 \\ \text{NUM} & \text{sg} \end{array}\right] \\
\text{OBJ} & \left[\begin{array}{ll} \text{PRED} & \text{`PRO'} \\ \text{PRON-TYPE} & \text{personal} \\ \text{CASE} & \text{acc} \\ \text{PERS} & 1 \\ \text{NUM} & \text{sg} \end{array}\right] \\
\text{TENSE} & \text{present} \\
\text{MOOD} & \text{indicative} \\
\text{VFORM} & \text{fin} \\
\text{CLAUSE-TYPE} & \text{interrog}
\end{array}\right]$$

The finite verb appears before a pronominal subject and so must be placed under `C`. Consequently, the clause is typed as interrogative by rule (32). The subject pronoun must be placed in the specifier of `IP` and the object pronoun is placed within the `VP` at the current stage of my model. Hence `I′` must re-write directly to `VP`, which is an option only if the clause type is not declarative. Since the clause type feature is specified as interrogative, this constraint is met. The model produces one solution for the parse.

(2) **Operator adverbs**. The second context that triggers placement of the verb under `C` involves special clause-initial operator adverbs. The most common items of this group are *þa*, (36a), and *þonne* 'then,' (36b). I will limit my discussion to those two words. However, other adverbs such as *swa* 'so', *nu* 'now' or *her* 'here, in this year' can sporadically license the verb under `C` as well.

(36) a. [$_{\texttt{CP}}$ **þa** *cwoman* [$_{\texttt{IP}}$ we to sumre byrig ]]
then came we to some city
'Then we arrived in a city' (coalex,Alex:15.1.126)

b. [$_{\texttt{CP}}$ **þonne** *findst* [$_{\texttt{IP}}$ þu ðæron ænne gyldenne wecg ]]
then find you thereon one golden coin
'Then you will find a golden coin in it' (cocathom1,ÆCHom_I,_34:470.163.6811)

The adverbs *þa* and *þonne* attract the verb to `C` almost categorically during the Old English period. It is only later, during the Middle English era, that 'subject - verb' order becomes common after *then* (for some early, Old English examples, see e.g., Mitchell 1985: §3922). Thus, my model must force the occurrence of the finite verb under `C` after these operator adverbs in the Old English period, but allow for a simple mechanism to lose attraction to the high verb position in subsequent periods at the same time.

The clause-initial adverb must have some special property that accounts for the fact that it attracts the verb to the high head position. Most scholars postulate a syntactically active feature for that purpose. For the goals of this study, it is irrelevant what this operator feature should be exactly. However, it would seem desirable to label the feature in a way that reflects its semantic contribution to some degree. The most promising approach to motivate verb placement under `C` after *þa* and *þonne* may be by reference to the idea that these items are rather special grammatical words used to create narrative cohesion. The words *þa* and *þonne* have been said to mark episodes (Foster 1975), function as additive and transitional connectives (Lenker 2010), signal specific kinds of continuity in a narrative (Wårvik 2011) or form a part of a system that manages discourse linking or anchoring (Los and Dreschler 2012). Following this line of thought, I will label the operator feature on *þa* and *þonne* 'presentational link,' `PRES-LINK=+`, alluding to the special way in which these adverbs present the proposition encoded in the main clause and in which they link it to the preceding discourse material. Example (37) shows the resultant lexicon entry for the operator adverb *þonne*.

(37) *þonne* ADV (↑ PRED)= 'THEN'
(↑ PRES-LINK)= +

A designated rule will regulate the introduction of operator adverbs into the clausal structure. Specifically, the rule will allow an adverb phrase, `ADVP`, with a presentational link feature to be placed in the specifier of $\mathtt{CP}_{MAT}$, where it will be interpreted as a topic in an underspecified discourse function, `TOP-UDF`, and unify with a member of the local adjunct set. I will make use of a special clause type for clauses that are declarative but also involve verb placement under `C`, called non-canonical declaratives, `CLAUSE-TYPE=non-can-decl`. When the operator adverb rule is employed, it automatically introduces such a non-canonical declarative clause type feature into the functional structure. This is shown in (38).

| (38) $\mathtt{CP}_{MAT} \rightarrow$ | ADVP | $\mathtt{C'}_{MAT}$ |
|---|---|---|
| | (↑ TOP-UDF) = ↓ | ↑ = ↓ |
| | (↑ TOP-UDF) ∈ (↑ ADJUNCT) | (↓ TENSE) |
| | (↓ PRES-LINK) $=_c$ + | |
| | (↑ CLAUSE-TYPE) = non-can-decl | |

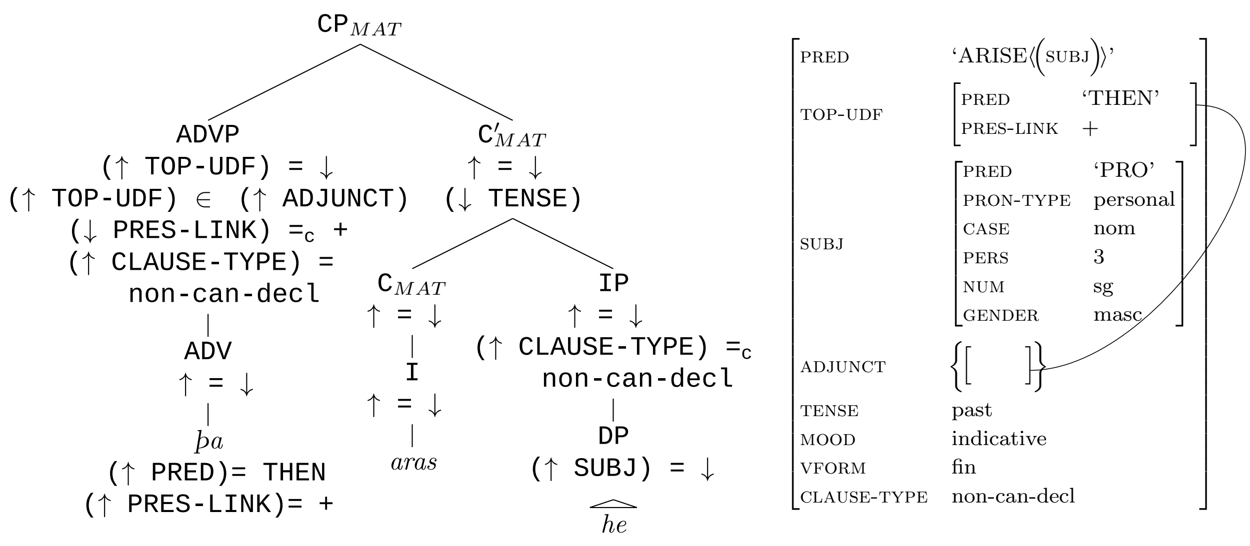
Old English operator adverbs are thus assumed to function largely identically to Modern English clause-initial negative constituents triggering subject-auxiliary inversion, like *never*, as discussed around example (4) in Appendix A. The main differences between, say, *þa* and *never* are the labels of the syntactically active feature, a presentational link feature in the former, a negative polarity feature in the latter case, and the degree to which the semantic repercussions of verb placement under `C` are currently understood and can be formalized.

The `IP`-layer will check for the presence of a non-canonical declarative clause type feature and allow the projection of a `C`-head if it is encountered.

| (39) $\mathtt{C'}_{MAT} \rightarrow$ | $\mathtt{C}_{MAT}$ | IP |
|---|---|---|
| | ↑ = ↓ | ↑ = ↓ |
| | | (↑ CLAUSE-TYPE) $=_c$ non-can-decl |

The tree in (40) presents an example parse of a sentences with clause-initial *þa* or *þonne*, followed by the finite verb and the subject. The sentence reads *þa aras he* 'Then he arose.'

(40) þa aras he
then arose he
'Then he arose' (cowsgosp,Mk_[WSCp]:2.14.2301)



The clause-initial adverb *þa* introduces a presentational-link feature into the structure. The rule that places the adverb in the specifier of CP, (38), is sensitive to this feature and thus values the clause-type attribute as a non-canonical declarative. This clause-type feature in turn is only compatible with re-writing C′ to a head and IP. Hence, the C-head must be projected; the finite verb must be placed under it, and will thus immediately follow the operator adverb, as desired. The model produces exactly one solution for the parse.

As long as *þa* and *þonne* are endowed with a presentational link feature, they can only be parsed by rule (38) so that the non-declarative clause type feature necessarily leads to the projection of a C-head by rule (39) and thus high verb placement. This accounts for the nearly obligatory nature of verb placement under C after operator adverbs in Old English.

The loss of high verb placement after *then* can be modeled as follows. I assume that adverb phrases are ordinarily parsed outside of the core clausal structure, as adjuncts to $\mathtt{CP}_{MAT}$. The relevant rule is shown in (41). It maps the initial ADVP onto an arbitrarily deeply embedded adjunct function, {XCOMP|COMP}$^*$ ADJUNCT, and can be applied only if its head is neither a *wh*-item nor an operator adverb.

(41) $\mathtt{CP}_{MAT} \rightarrow$

| ADVP | $\mathtt{CP}_{MAT}$ |
|---|---|
| $\downarrow \in$ ($\uparrow$ {XCOMP\|COMP}$^*$ ADJUNCT) | $\uparrow = \downarrow$ |
| ($\downarrow$ ADV-TYPE) $\neq$ wh | |
| $\neg$($\downarrow$ PRES-LINK)= + | |

An adjunction rule of this kind is required by the fact that adverbs can occur in front of constituents assumed to be in the specifier of CP. These cases include adverbs before topicalized nominal arguments, (42a), before *wh*-items in direct questions, (42b), or before operator adverbs, (42c). The examples illustrate with the adverb *soþlice* 'truly', often encountered in biblical writings. A similar adjunction rule to CP has also been proposed for scene-setting adverbials, including PPs (Kroch and Taylor 1997: 304). An example is shown in (42d).

(42) a. **Soðlice** [$_{\mathtt{CP}}$ Drihtne [$_{\mathtt{IP}}$ þu *agylst* þine aðas ]]
truly to-God you repay your oaths
'Truly, you fulfill your oaths to God' (cowsgosp,Mt_[WSCp]:5.33.261)

b. **Soþlice** [$_{\mathtt{CP}}$ hwam *telle* [$_{\mathtt{IP}}$ ic þas cneorysse gelice ? ]]
truly to-whom tell I this generation like
'Truly, what can I compare this generation to?' (cowsgosp,Mt_[WSCp]:11.16.679)

c. And **soðlice** [$_{\mathtt{CP}}$ þa *com* [$_{\mathtt{IP}}$ stefn of heofenum ]]
and truly then came voice of heaven
'And truly, then a voice came from heaven' (cowsgosp,Mt_[WSCp]:3.17.159)

d. **On þa ilcan tima** [$_{\mathtt{CP}}$ þa *comon* [$_{\mathtt{IP}}$ hi to Medeshamstede ]]
in the same time then came they to Medhamsted
'At the same time, they came to Medhamsted' (ChronE_[Plummer]:870.5.1115)

All that is now required is that English develops an alternative lexicon entry for *then* that does not have special presentational or discourse-linking properties, hence does not include a presentational link feature. Such an innovative lexicon entry for *then* is shown in (43).

(43) *þonne* ADV ($\uparrow$ PRED)= 'THEN'

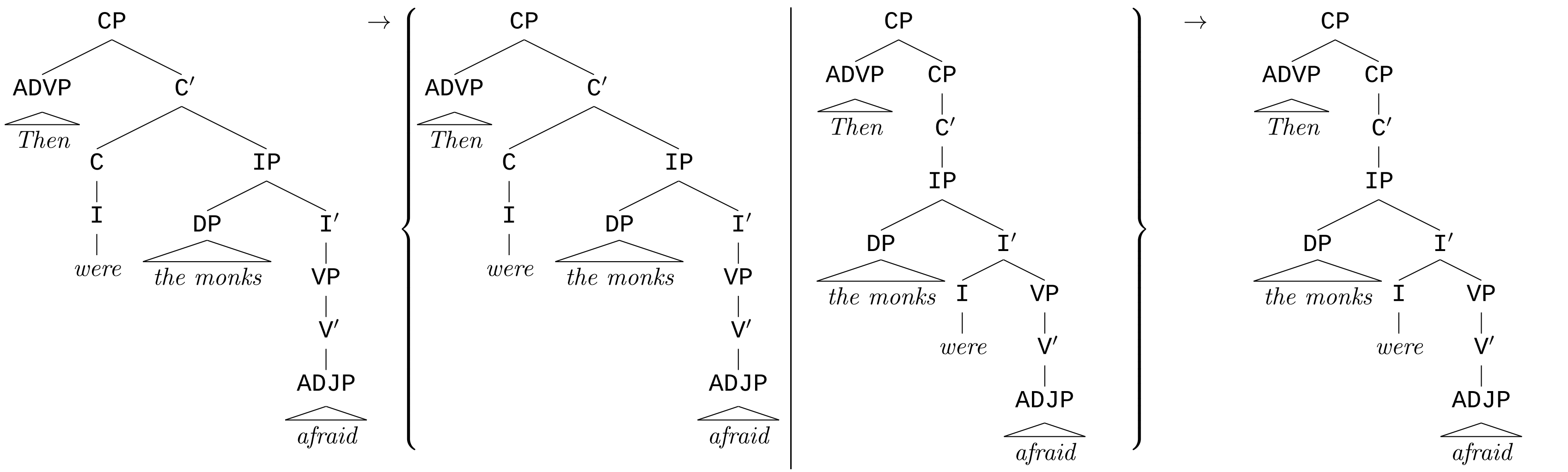
Consequently, initial innovative *then* can only be parsed with the rule in (41) and will thus not value the clause type as a non-canonical declarative or trigger verb placement under C anymore. Below, I cite an example from the late 14$^{\text{th}}$ century, when 'subject - verb' order after *then* has become the norm.

(44) [$_{\mathtt{CP}}$ **Thanne** [$_{\mathtt{CP}}$ Job preiede ]]
then Job prayed
'Then, Job prayed' (CMPURVEY,I,37.1747) (c. 1388 A.D.)

By letting the conservative and innovative two lexicon entries for *then* in (37) and (43) compete and the former gradually drive out the latter, one can thus account for the eventual disappearance of V-to-C structures after operators adverbs. The change is illustrated in the graphic below, which, once again, presents the development in terms of the general model of the implementation of a linguistic innovation in a group of speakers from the introductory chapter.

**Change in Verb Placement: Loss of V-to-C after Operator Adverbs**

| Conservative period: Speakers represent only the conservative variant $\alpha$ | Transitional period: Speakers represent both the conservative variant $\alpha$ and the innovative variant $\beta$ | Innovative period: Speakers represent only the innovative variant $\beta$ |
|---|---|---|
| $\alpha$ $\rightarrow$ | $\{$ $\alpha$ $\mid$ $\beta$ $\}$ | $\rightarrow$ $\beta$ |
| Conservative period: Speakers represent *then* only as operator adverb | Transitional period: Speakers represent *then* both as operator and as an ordinary adverb | Innovative period: Speakers represent *then* only as an ordinary adverb |



(3) **Negation**. Finite verbs are also licensed under `C` if they are negative. Sentential negation in Old English is achieved by cliticizing the negative particle *ne* to the left of the finite verb, (45a). Negation can also be fused to the verbal stem, as in *nis* = *ne is* 'is not', *nolde* = *ne wolde* 'would not' etc, (45b).

(45) a. [$_{\mathtt{CP}}$ **Ne** *geseah* [$_{\mathtt{IP}}$ he Crist on life ]]
not saw he Christ in life

'He did not see Christ alive' (coaelive,ÆLS_[Mark]:147.3294)

b. [$_{\mathtt{CP}}$ **n***yton* [$_{\mathtt{IP}}$ hi hwæt hi doð ]]
not-know they what they do

'They don't know what they do' (cocathom2,ÆCHom_II,_2:17.183.408)

Sentential negation can license verbs under `C`, as in the examples above, but does not categorically force them into this position. That is to say, negated verbs can also appear after their subjects in `I`. It is difficult to discern semantic differences between these two positions. The sentences in (46) illustrate negated verbs in `I` with nearly synonymous meanings to the above examples of negated verbs in `C`.

(46) a. [$_{\mathtt{CP}}$ [$_{\mathtt{IP}}$ He **ne** *geseah* hine siððan ]]
He not saw him then

'He did not see him then' (coaelive,ÆLS_[Book_of_Kings]:296.3879)

b. [$_{\mathtt{CP}}$ [$_{\mathtt{IP}}$ ge **n***yton* hwænne seo tid ys ]]
you not-know when the time is

'You don't know when the time will come' (cowsgosp,Mk_[WSCp]:13.33.3255)

Thus, my model should make placement of negated verbs under `C` possible but not mandatory. I will not assume any systematic meaning differences between negated verbs under the higher and lower head position.

Sentential negation will be modeled as follows. I will use the part of speech label `NEG` for the word class of the negative particle *ne* and the compound label `NEG+I` for negatively contracted forms such as *nis* 'is not.' These categories introduce a negative polarity feature, `POL=neg`. The lexicon entry for the negative particle below illustrates. Forms with fused negation will contain the same feature.

(47) *ne* `NEG (↑ POL)= neg`

Next, an alternative phrase structure rule for `C`$_{MAT}$ is required. It states that if the domain of the `C`-head includes a negative polarity feature, then the clause type will be valued as a non-canonical declarative. Verb placement under `C` can then automatically be regulated by the same clause type mechanism as for the previous operator adverb context.

(48) `C`$_{MAT}$ →
```
            I
          ↑ = ↓
      (↓ POL)=c neg
 (↑ CLAUSE-TYPE)= non-can-decl
```

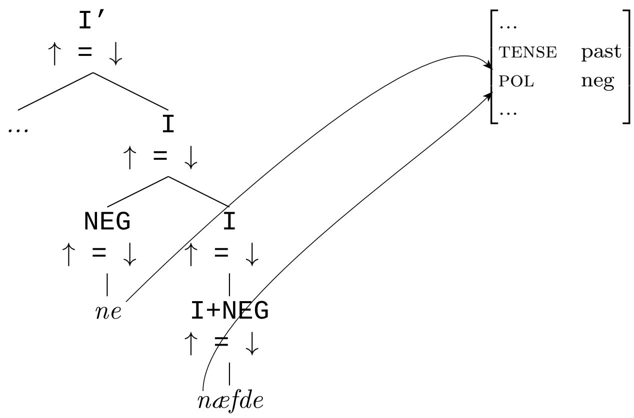
Finally, I assume that the head `I` can re-write to negation and the finite verb, (49a), or to a form that shows negation contracted to the verbal stem, (49b). The introduction of negation is thus modeled as head adjunction.

(49) a. `I` →
```
   NEG          I
  ↑ = ↓       ↑ = ↓
```
b. `I` →
```
   NEG+I
   ↑ = ↓
```

The model developed here leads to three noteworthy consequences. First, the re-write rules for `I` in (49) are as applicable under `C` in non-canonical declaratives as they are under `I` in ordinary declaratives. Hence, the grammar can handle clauses with negated verbs in front of subjects as well as in post-subject position. This successfully models the desired optional nature of placement under `C` for negated verbs.

Secondly, the model allows, in principle, for multiple realizations of sentential negators. The reason is that the rule in (49a) is self-feeding and that the polarity features introduced by several instances of negation are all identical and can thus unify. Consequently, any number of sentential negators would be expected to value the polarity of a sentence as negative. Given that Old English generally displays negative concord behavior, this may not be an implausible corollary. Old English is different in this respect than its modern successor, as discussed for Modern English in chapter 2 around example (34). In addition, I did in fact find one Old English example that seems to appear with two sentential negations, one free, the other contracted to the verbal stem. My model can assign a structure to such patterns, as shown below.

(50) ... þæt he from his abbode ne onfeng and þurh his sylen **ne n**æfde
... what he from his abbot not received and through his self not not-had
'... that which he did not receive from his abbot and did not have by himself'
(cobenrul,BenR:55.91.18.1008)



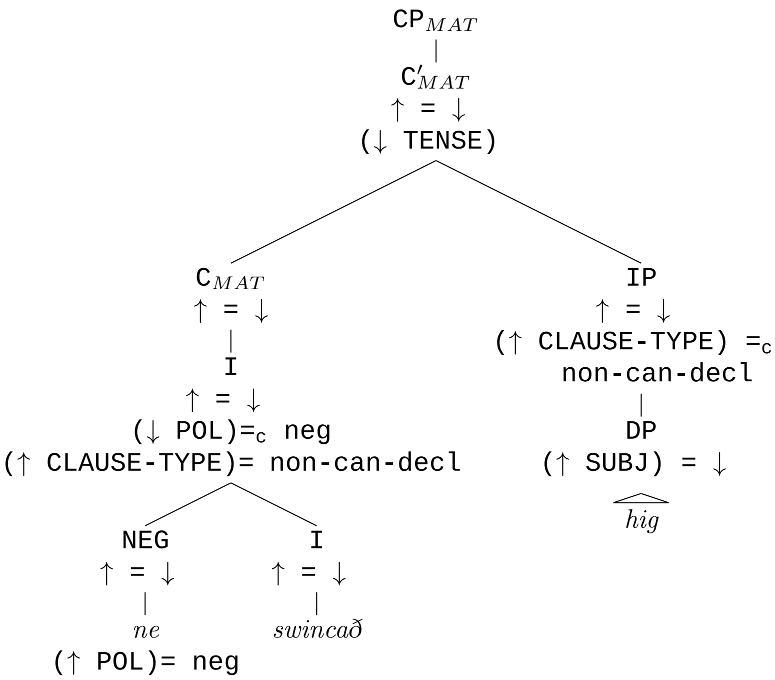

Thirdly, the model will produce two solutions for parses of negative verb-first sentences. This kind of ambiguity does not count as an unwanted byproduct of my particular grammar implementation, but reflects, I believe, an actual characteristic of Old English syntax. Negative verb-first sentences systematically allow for two distinct readings. (i) They can be non-canonical declaratives. This interpretation arises from placement of the verb under C by rule (48), which introduces a non-canonical clause type feature and thus triggers the projection of a C-head by rule (39). (ii) They can be interrogatives. This interpretation is produced by placing the verb under C with the rule in (33), which does not lead to any valuation of the clause type attribute. As a consequence, the only parsing strategy arriving at a consistent constituent structure must employ the rule in (32), resulting in the introduction of interrogative force. For example, the two sentences in (45) could very likely also mean 'Didn't he see Christ alive?' and 'Don't they know what they do?' respectively. Examples of attested questions with similar structures are shown in (51) below. The second example shows an eroded ending on the verb.

(51) a. ac [$_{\text{CP}}$ **ne** geseah [$_{\text{IP}}$ ic þe on ðam wyrtune mid him? ]]
but not saw I you in the garden with him
'But didn't I see you with him in the garden?' (coverhomE,HomS_24.1_[Scragg]:173.147)

oððe [$_{\text{CP}}$ **n**yte [$_{\text{IP}}$ ge la þæs sceoccan deopnysse? ]]
or not-know you lo of-the fiend's deepness
'Or don't you know, verily, the fiend's cunningness?' (coaelhom,ÆHom_30:110.4131)

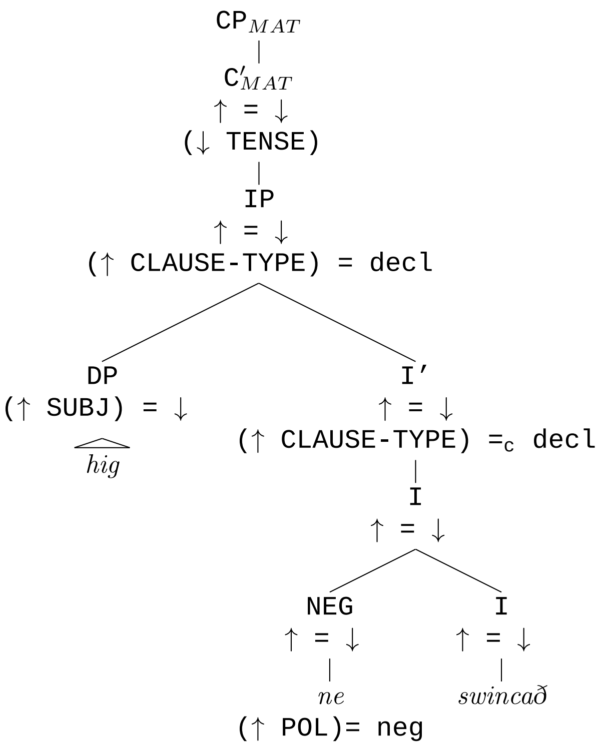
My grammar fragment can now parse declaratives with negative verbs licensed in the high or the low head position. An example is provided by the attested minimal pair *ne swincað hig* and *hi ne swincað*, both meaning 'they do not toil.'

(52) ne swincað hig
not toil they
'They do not toil' (cowsgosp,Mt_[WSCp]:6.28.339)




| | | | |
|---|---|---|---|
| PRED | 'TOIL⟨(SUBJ)⟩' | | |
| SUBJ | | PRED | 'PRO' |
| | | PRON-TYPE | personal |
| | | CASE | nom |
| | | PERS | 3 |
| | | NUM | pl |
| TENSE | present | | |
| MOOD | indicative | | |
| VFORM | fin | | |
| POL | neg | | |
| CLAUSE-TYPE | non-can-decl | | |

(53) hi ne swincað
they not toil
'They do not toil' (cowsgosp,Lk_[WSCp]:12.27.4686)




| | | | |
|---|---|---|---|
| PRED | 'TOIL⟨(SUBJ)⟩' | | |
| SUBJ | | PRED | 'PRO' |
| | | PRON-TYPE | personal |
| | | CASE | nom |
| | | PERS | 3 |
| | | NUM | pl |
| TENSE | present | | |
| MOOD | indicative | | |
| VFORM | fin | | |
| POL | neg | | |
| CLAUSE-TYPE | decl | | |

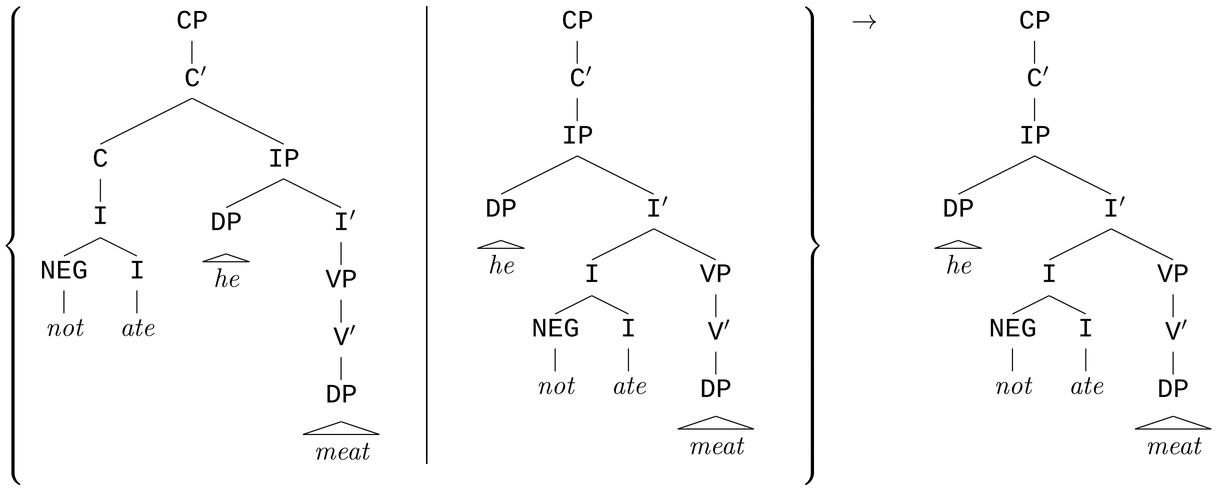
In these examples, the negative particle *ne* is adjoined to the finite verb and values the polarity of the sentence as negative by rule (49a). In the first sentence, the negative polarity feature allows parsing the negation+verb cluster under `C` in accordance with rule (48). This has the effect of typing the clause as a non-canonical declarative. The non-canonical clause type feature, in turn, enables the overt `C`-head to occur with `IP` as its complement, leading to a `V`-to-`C` structure. The grammar unifies all features and produces the solution for the parse shown above. (As mentioned earlier, a second solution can be produced by my model, which interprets the sentence as an interrogative clause, meaning 'Do they not toil?', but this reading is not relevant here and hence not displayed.) In the second sentence, the matrix `CP`-layer does not project an overt head, the clause is thus interpreted as a declarative, and the verb must consequently be placed under `I`. The model generates one unique solution for the sentence.

I interpret the variation between high and low negated verbs as an indication of a change in progress. The rule licensing negated verbs under `C` in (48) is gradually lost. Unfortunately, the textual record of Old English does not stretch back far enough to determine with certainty if conservative high verb placement with negative polarity ever was mandatory. Variation between high and low negated verbs exists from the earliest Old English documents on. The figure below illustrates the development of negative `V`-to-`C` in terms of the general model of competition between grammatical options outlined in the introduction chapter. I leave out an illustration of the conservative period and merely indicate with a question mark the fact that the existence of a pure conservative state is uncertain.

**Change in Verb Placement: Loss of `V`-to-`C` under Negation**

| | | |
|---|---|---|
| Conservative period: Speakers represent only the conservative variant $\alpha$ | Transitional period: Speakers represent both the conservative variant $\alpha$ and the innovative variant $\beta$ | Innovative period: Speakers represent only the innovative variant $\beta$ |
| $\alpha \rightarrow$ | $\{ \ \alpha \ \mid \ \beta \ \} \rightarrow$ | $\beta$ |
| Conservative period: Unknown if speakers ever obligatorily required high verb placement under negation | Transitional period: Speakers represent both a specific rule licensing `V`-to-`C` with negation, and a general rule for verb placement under `I` that can apply to negated verbs | Innovative period: Speakers represent only a general rule for verb placement under `I` that can apply to negated verbs |

?



(4) **Non-indicative mood**. The fourth context that allows Old English finite verbs to occur under `C` involves subjunctive and imperative mood. In matrix clauses, the two non-indicative moods seem to form one relatively cohesive system dealing with wishes, encouragement and commands. The attenuation of inflectional endings before and during the Old English period may have contributed to the coalescence of subjunctive and imperative mood to form such a unified 'imperative-hortative system' (Auwera et al. 2005). For instance, present singular subjunctives in *-e*, e.g., *þence* 'would think,' can easily be confused with singular imperatives without an ending, e.g., *þenc* 'think!'. In fact, Old English already has numerous forms that are syncretic between imperative and subjunctive mood, such *beo* 'would be / be!'. Main clause subjunctives in the third person express an optative, a realizable wish, (54a), in the first person plural an adhortative, an encouragement including the speaker (54b). There are very few examples of second person subjunctives expressing wishes or exhortations (Mitchell 1985: §892). Instead, the imperative supplies these forms, (54c). Imperatives do not necessarily occur with overt subjects. The realization of an overt subject is more likely if the imperative is also negated, as in the example below.

(54) a. [$_{CP}$ **Nime** [$_{IP}$ he þisne clað ]]
take.subjunctive he this cloth
'{May he/ Let him / He should } take this cloth' (cocathom2,ÆCHom_II,_31-32:242.34.5375)

b. Men, [$_{CP}$ **geearnian** [$_{IP}$ we nu þæt ure se ytmesta dæg sy engla gefea ]]
men earn.subjunctive we now that our the last day be angels' joy
'Men, {let us / we should } now earn it that our last day be the joy of angels'
(coblick,HomU_19_[BlHom_8]:101.95.1308)

c. [$_{CP}$ **ne synga** [$_{IP}$ þu heononforð ]]
not sin.imperative you henceforth
'Do not sin anymore!' (cocathom1,ÆCHom_I,_24:378.204.4790)

As a side note, the subjunctive mood's matrix clause function to wish, urge and command may be related to its etymological origin. The Germanic subjunctive derives from the Proto-Indo-European optative. For example, the present, active, optative third person singular of Proto-Indo-European ***$b^h$*er-* 'carry,' **$b^h$*éroyd* (Ringe 2006: 39) yields the Ancient Greek optative *φέροι* 'may he carry / if only he could carry,' but also the Proto-Germanic subjunctive **berai* (ibid: 237), as in Gothic *bairai* or Old English *bere* 'he might / should / would carry.' The old optative usually expressed wishes and potentialities in main clauses. Hence, it seems plausible that this core function of the mood has been preserved in Old English main clauses.

Verb placement under `C` should be constructed as optional in the non-indicative environment, comparable to the negation context. This becomes obvious from the fact that subjunctives and, probably less commonly, imperatives can appear after the subject in `I` as well. As an illustration, the examples with non-indicative verbs before subjects in (54) can be compared to the identical verb forms after subjects in (55).

(55) a. [$_{IP}$ he **nime** lynen hrægl ]
he take.subjunctive linen cloth
'{May he/ Let him / He should } take a linen cloth' (colwstan2,ÆLet_3_[Wulfstan_2]:68.79)

b. ac [$_{IP}$ we **geearnien** þæt we moten [...] becuman to þære ecan eadignesse ]
but we earn.subjunctive that we may [...] come to the eternal blessedness
'But {let us / we should } earn it that we will come [...] to the eternal blessedness'
(coverhomL,HomU_15.1_[Scragg]:178.94)

c. and [$_{IP}$ þu heonan forð **ne synga** ]
and you hence forth not sin.imperative
'Do not sin anymore!' (coaelhom,ÆHom_14:226.2127)

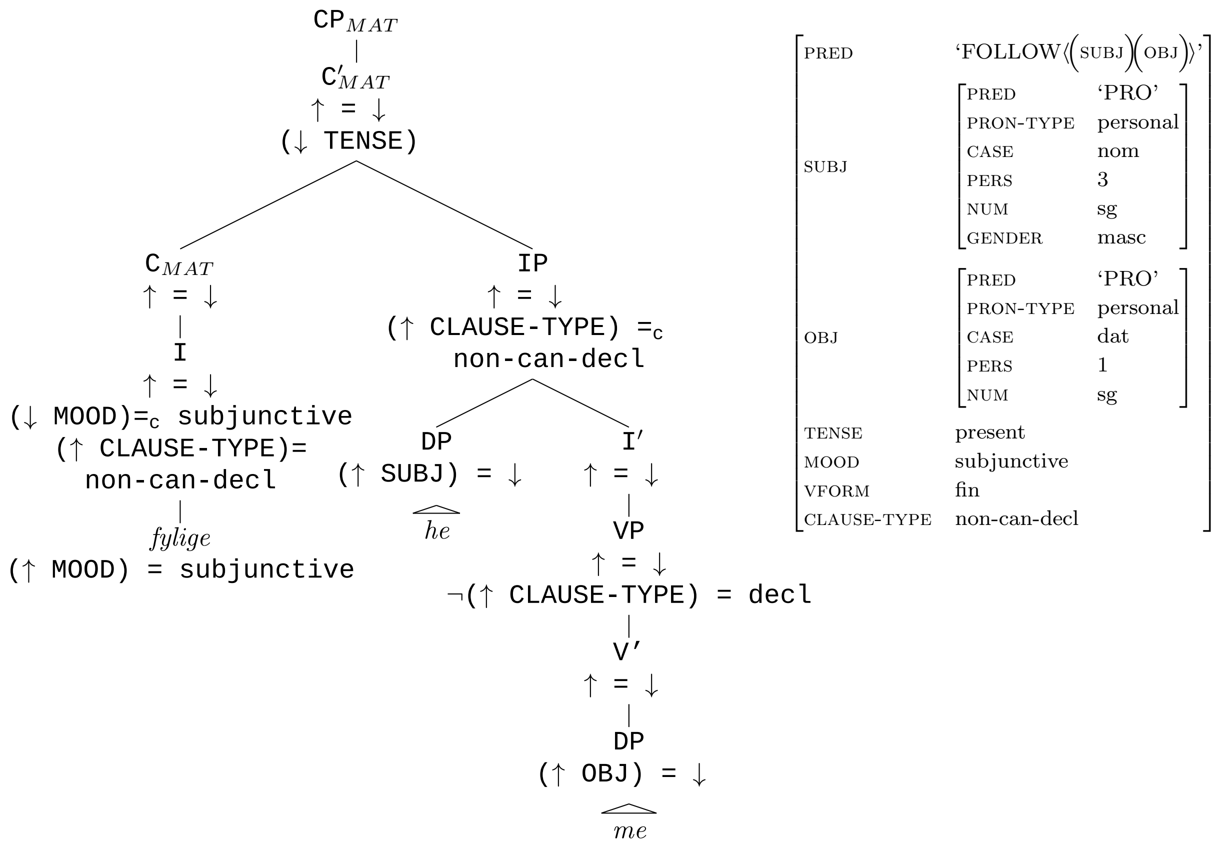
It is important to note that the inflectional endings that mark subjunctive vs. indicative mood are in the process of eroding away by the time Old English is written down. One can often not be sure if scribes really intended the subjunctive suggested by an apparently subjunctive ending or if they weakened the ending of an indicative. When one reads *Hi eoden*, the formal appearance of a subjunctive in *-en* 'they should go' seems as likely as a weakened indicative ending *-on>en* 'they went.' Many forms are already ambiguous between indicative and subjunctive mood in Old English, such as *þu sawe*, which can mean both 'you may see' or 'you saw.' Therefore, it becomes difficult to systematically distinguish subjunctives from indicatives on the basis of their formal realization. As a consequence, the non-indicative mood trigger for verb placement under `C` is quite difficult to identify and measure in corpus data.

I will now implement a mechanism for the optional placement of imperatives and subjunctives under `C`. The addition of the single rule in (56) will suffice to accomplish this goal. It allows the `C`-node to re-write to a finite verb of category `I` if it is subjunctive or imperative and, if this is the case, introduces a non-canonical-declarative clause type feature. This feature will then regulate verb placement under `C` as before.

(56) $\mathtt{C}_{MAT} \rightarrow$
```
                    I
                  ↑ = ↓
          { (↓ MOOD)=c subjunctive
          | (↓ MOOD)=c imperative }
        (↑ CLAUSE-TYPE)= non-can-decl
```

My model can now parse non-indicative moods under `C` as well as `I`. I will illustrate for a subjunctive hortative in pre-subject position, *fylige he me* 'may he follow me / let him follow me / he should follow me.' Non-indicative verbs following subjects will be parsed with the rules for ordinary declarative clauses placing the verb in `I` as in the negation context above.

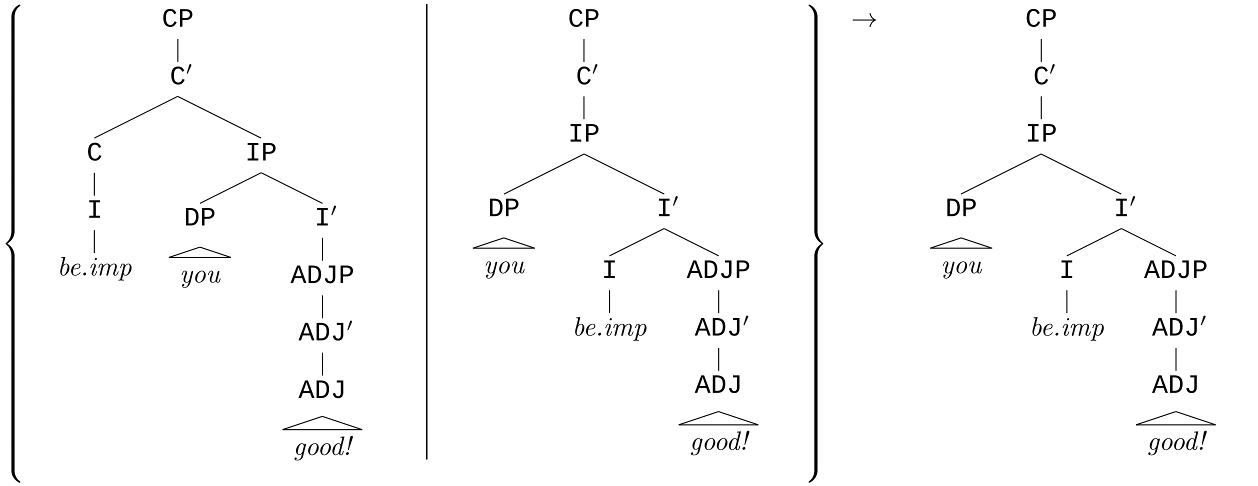
(57) **fylige** he me.
follow.subjunctive he me
'{May he / Let him / He should } follow me.' (cocathom1,ÆCHom_I,_10:263.141.1930)




Rule (56) can place the subjunctive verb under `C` because of the appropriate mood licensing constraints. The clause type is thus valued as a non-canonical declarative. The clause-type sensitive rule in (39) can therefore be employed to project `IP` as the complement of the `C`-head. The model produces one unique solution for the analysis of this sentence.

The variation between high and low non-indicative verbs is also assumed to be due to a change in progress. The rule licensing high subjunctives or imperatives in (56) is lost over time. Variation between hortatives parsed with `V`-to-`C` or the verb within the `IP` exists from the earliest Old English records so that one cannot know with certainty if the former structure ever was obligatory. Furthermore, the Old English mood distinctions erode away so that the licensing conditions for the syntactic rule in (56) fade out of the language before the change is complete. The figure shown below presents the development as a competitive change according to the general model of linguistic changes from the introductory chapter.

**Change in Verb Placement: Loss of V-to-C with Non-Indicative Mood**

| Conservative period: Speakers represent only the conservative variant $\alpha$ | Transitional period: Speakers represent both the conservative variant $\alpha$ and the innovative variant $\beta$ | Innovative period: Speakers represent only the innovative variant $\beta$ |
|---|---|---|
| $\alpha \rightarrow$ | $\{ \alpha \mid \beta \} \rightarrow$ | $\beta$ |
| Conservative period: Unknown if speakers ever obligatorily required high verb placement with hortatives | Transitional period: Speakers represent both a specific rule licensing V-to-C with hortatives, and a general rule for verb placement under I that can apply to hortative verbs | Innovative period: Speakers represent only a general rule for verb placement under I that can apply to hortative verbs (formally marked subjunctives / imperatives with overt subjects soon disappear from the system) |
| ? |  | |

Before I finish this section, I would like to discuss exceptions to the generalizations discussed so far, i.e., verbs that occur before pronominal subjects but are positive, indicative in non-interrogative main clauses without operator adverbs. I will not amend my model in an attempt to enable the generation of such structures. The reasons for this decision are as follows:
First of all, such patterns are very rare. A model that cannot analyze such sentences thus loses little empirical coverage.
Secondly, indicative verb-first main clauses occur substantially more frequently in some Old English texts, such as *Bede*, than in others. This indicates idiosyncratic preferences among Old English authors as well as potential Latin interference effects. An example from *Bede* is shown below.

(58) [$_{\mathsf{CP}}$ *Comon* [$_{\mathsf{IP}}$ hi of þrim folcum, ðam strangestan Germanie ]]
came they from three peoples the strongest of-Germania

‘They came from the three strongest peoples of Germania’ (cobede,Bede_1:12.52.2.469)

The Latin original shows the finite verb in first place. It is therefore possible that the translator was influenced by this word order and mechanically added a subject pronoun after the verb.

(59) *Latin*

*Advenerant* autem de tribus Germaniæ populis fortioribus
had-came however from three of-Germania peoples strong

‘They had come, however, from three strong peoples of Germania’
(Moberly 1881: 36, book 1, chapter 15)

Thus a model that allows for the unrestricted analysis of indicatives under `C` is likely to misconstrue the reality of Old English prosaic syntax.
Thirdly, where indicatives do appear before subject pronouns in untranslated Anglo-Saxon texts, they seem to be linked to a rather special narrative style. In particular, they seem to express dramatic and lively episodes. They may also be close to the spoken word as suggested by comparisons with modern languages such as Dutch (e.g., Kemenade 1987: 44-5) and the fact that they usually occur with copulas or pre-modals. An example is shown in (60)

(60) [$_{\mathsf{CP}}$ *Wæron* [$_{\mathsf{IP}}$ hi eac swyþe druncene ]]
were they also very drunk

‘They were also very drunk’ (ChronC_[Rositzke]:1012.8.1536)

The application of the rule licensing such structures would thus be functionally determined by a particularly idiosyncratic setting, a kind of narrative, bardic register. Since my model is supposed to capture core syntactic phenomena, I decided to leave such structures unconsidered.
Fourth, there may be other, lexically restricted idioms that allow indicative verbs to occur under `C`. The most important one seems to be the quotatives *cweþan* ‘say, tell’ and *secgan* ‘say, declare.’ The following example illustrates.

(61) Ic lufige God ælmihtigne, [$_{\mathsf{CP}}$ *cwæð* [$_{\mathsf{IP}}$ hi ]]
I love God almighty quoth she

‘ “I love God almighty,” she said.’ (comargaC,LS_14_[MargaretCCCC_303]:6.8.77)

The option for high verb placement in this lexically determined construction has survived to the present day. One can still encounter *said he*, or archaic *quoth he* etc. It would be possible to implement such form-meaning correspondences that are larger than individual words in LFG without admitting constructions in a technical sense as primitive entities in the ontology of the framework (Asudeh et al. 2013). However, the specific implementation and consequences of the technological extensions are still being debated (e.g., Müller and Wechsler 2014) and so complex that I find it unwarranted to include individual, lexicalized constructions such as quotatives into my toy grammar at this point.

This concludes my discussion of contexts triggering high verb placement in Old English.

#### 3.3.1.5 Extraposition

Old English permits extraposition of a wide variety of constituents. Postposed constituents can be identified as post-verbal elements in **I**-final structures. In the following examples, a pre-verbal nominal object forces **I**-final headedness. The post-verbal indirect object DP in (62a) must therefore have extraposed. Similarly, non-finite verbs as in (62b), or even entire verb phrases, can occur in configurations that are indicative of extraposition in Old English. The latter structures are usually referred to as " verb (projection) raising" in the literature.

(62) a. [$_{\text{IP}}$ ic oðerne oxan *geoffrige* ] [$_{\text{DP}}$ minum Gode ]
I other ox offer to-my god
'I will offer another ox to my God' (coaelive,ÆLS_[Book_of_Kings]:104.3727)

b. [$_{\text{IP}}$ ic þa ða wynstran dælas Indie *wolde* ] [$_{\text{V}}$ geondferan ]
I then the left parts of-India wanted through-go
'I then wanted to explore the left-hand regions of India' (coalex,Alex:26.7.312)

Although a core syntactic phenomenon of Old English grammar, extraposition is not crucial for this study and I will therefore deal with it rather superficially. The phenomenon is relevant here only in so far as the structural ambiguity it causes will lead to certain restrictions on data collections in the second, empirical part of this chapter. Furthermore, the phenomenon is still relatively poorly understood, which makes a proper implementation difficult. For instance, it is not clear where the extraposed constituent should be adjoined. It is also unknown if or to what degree there is a change in the frequencies of postposed constituents during the Early English period. Most importantly, the functional conditions exploiting the grammatical option of extraposition are not well understood at the moment. It is likely that postposition serves a number of interrelated functions, perhaps to express focus or to facilitate the processing of phonologically heavy elements.

I will model extraposition as adjunction to IP. Further, my model will map material dominated by the extraposed node onto an extraposed underspecified discourse function, EP-UDF. Not all sentence elements can postpose in Old English. In particular, pronouns, particles, stranded prepositions, and negatively quantified objects cannot occur in extraposed configurations. I will model these restrictions in part as the unavailability of extraposition phrase structure rules for certain categories, e.g., the lack of such a rule for particles, and in part as syntactic constraints on extraposition rules, e.g., constraints banning pronominal heads or negative quantification. For brevity's sake, the rule in (63) exemplifies the licensing conditions on extraposition for only one of many postposing categories. Specifically, the rule allows direct objects, OBJ, or indirect objects, $\text{OBJ}_{th}$, to occur to the right of IP. They will then receive some particular information-structural interpretation by virtue of being unified with an EP-UDF, which is not modeled here in any detail. The rule can only apply if the object is not headed by a pronoun and if it is not negatively quantified. Ultimately, these constraints could probably be expressed more elegantly in terms of information structure and a licensing mechanism for negative concord. It is even conceivable that what is currently conceptualized as extraposition could be analyzed in a more comprehensive way altogether. However, the current formalization will suffice for my purposes.

(63) a.
$$\text{IP} \rightarrow \quad \begin{array}{c} \text{IP} \\ \uparrow = \downarrow \end{array} \quad \begin{array}{c} \text{DP} \\ (\uparrow \text{EP-UDF}) = \downarrow \\ \{(\uparrow \text{EP-UDF}) = (\uparrow \text{OBJ}) \\ \mid (\uparrow \text{EP-UDF}) = (\uparrow \text{OBJ}_{th})\} \\ (\downarrow \text{PRED}) \neq \text{PRO} \\ \neg(\downarrow \text{QUANT POL}) = \text{neg} \end{array}$$

As a consequence of the introduction of this rule, most 'subject - verb - object' structures will be ambiguous between an **I**-initial and a postposition parse. This is illustrated in (64) below (see also (26b)).

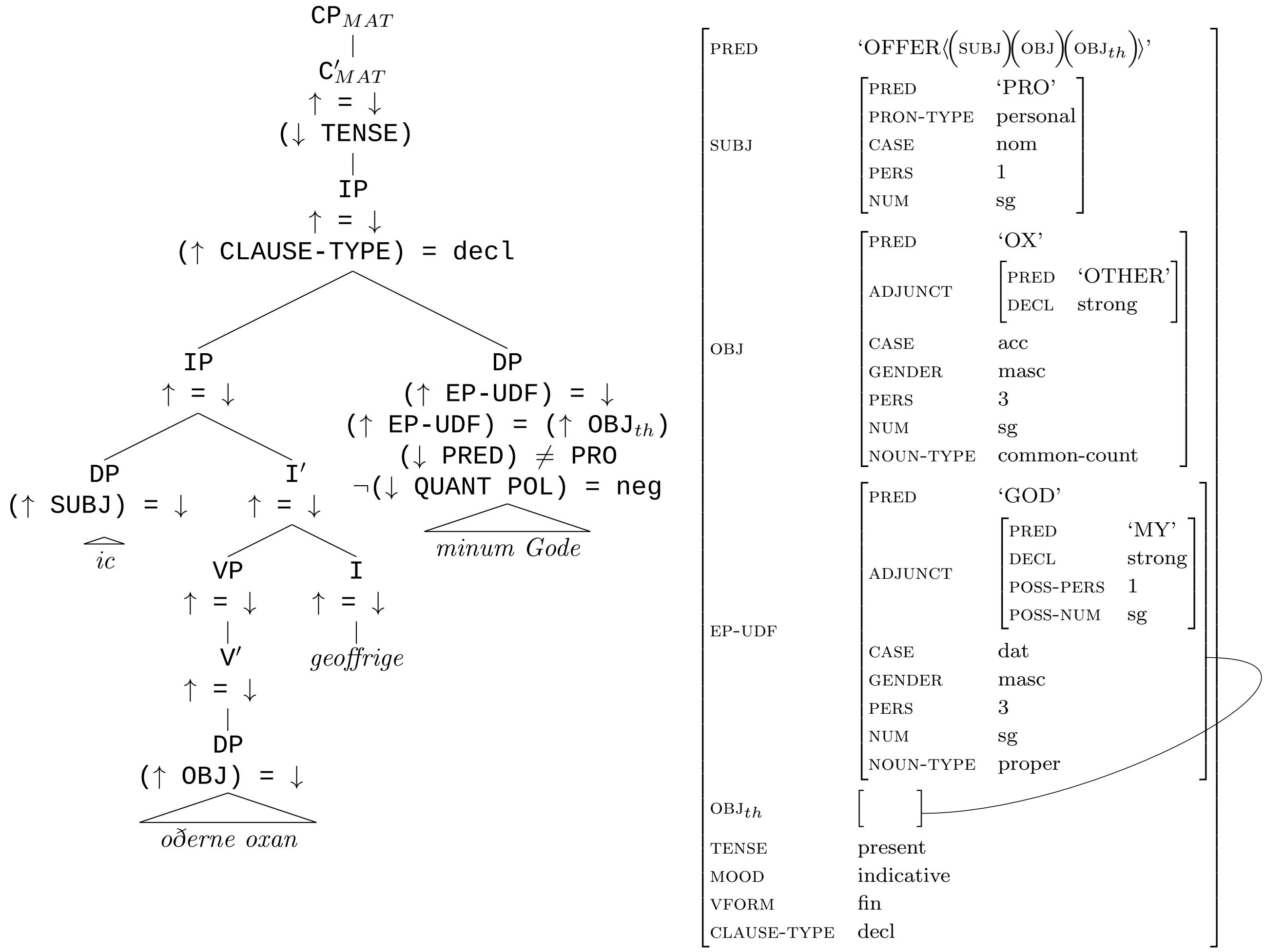
(64) a. Ic *can* þone dom
I know the doom
'I know the condition' (coapollo,ApT:4.12.45)

b. ***I**-Initial parse*
[$_{\text{IP}}$ Ic [$_{\text{I}'}$ *can* [$_{\text{VP}}$ [$_{\text{DP}}$ þone dom ] ]]]

c. *Extraposition parse*
[$_{\text{IP}}$ [$_{\text{IP}}$ Ic [$_{\text{I}'}$ *can* ]] [$_{\text{DP}}$ þone dom ] ]

My grammar model can now handle structures with postposed object arguments. As an example, I show my model parse for the sentence in (62a) in (65) below.

(65) ic oðerne oxan geoffrige minum Gode
I other ox offer to-my god
'I will offer another ox to my God' (coaelive,ÆLS_[Book_of_Kings]:104.3727)




The example is analyzed as follows. The subject pronoun must be placed in the specifier of IP as the only subject position for pronouns in my model. The following direct object occurs between subject and verb and can therefore only be parsed as an element inside the VP. The pre-verbal object forces I-final headedness. Consequently, post-verbal material, here an indirect object, can only be analyzed as an extraposed element. Specifically, the indirect object DP is right-adjoined to IP with the extraposition rule in (63). The rule can be applied here because the extraposed DP is neither pronominal nor negatively quantified. The model produces one solution for the parse (although technically two due to syncretism of the verb in indicative and subjunctive mood).

#### 3.3.1.6 Topicalization

It is not uncommon for prominent Old English constituents to be placed in a discourse function in initial position in main clauses, a phenomenon commonly referred to as 'topicalization.' By 'topic,' I only refer to the syntactic position of the fronted element, not its discourse function. I will model the clause-initial topic position as the specifier of CP. The following sentences illustrate. Example (66a) presents a fairly typical case of topicalization - a fronted demonstrative anaphor establishes a discourse link with a previously mentioned, given referent. Example (66b) shows a topicalized PP that could be described as a kind of attention focus spotlighting a particular aspect of the event described for narrative, dramatic or emphatic effect. Finally, example (66c) illustrates two topic DPs that appear to be contrastive. There are many other kinds of discursive roles that Old English topics can take, and an even larger number of conceptualizations of their ontology (see e.g., Meurman-Solin et al. (2012) for a collection of papers on information structure in the history of English, all with essentially different perspectives on analyzing discourse). My model will not attempt to account for the different information-structural interpretations of topics in any way.

(66) a. Æðelhere$_i$ Annan broðor Eastengla cyninges [...] se wæs ordfruma þæs gefeohtes.
Æthelhere Anna's brother of-East-Angles king he was point-origin of-the war
[$_{\text{CP}}$ [$_{\text{DP}}$ **þone$_i$** ] [$_{\text{IP}}$ mon slog eac mid ealle his compweorode ]]
that-one one slew also with all his battle-troop

'Æthelhere, the brother of Anna, the king of East-Anglia [...], he was the instigator of the war. That one, people killed along with all his troops'
(cobede,Bede_3:18.236.15.2411-2413)

b. [$_{\text{CP}}$ [$_{\text{PP}}$ **on weallendum ele** ] [$_{\text{IP}}$ he het hine baðian ]]
in boiling oil he ordered him bathe

'In boiling oil, he ordered that one should bathe him'
(colsigewZ,ÆLet_4_[SigeweardZ]:1026.455)

c. [$_{\text{CP}}$ [$_{\text{DP}}$ **þan synfullen** [...] ] [$_{\text{IP}}$ he behet Godes godnysse [...] ]],
to-the sinful he promised God's goodness
[$_{\text{CP}}$ [$_{\text{DP}}$ **þa gode** ] [$_{\text{IP}}$ he manode, þæt heo on heora godnysse þurhwunedan ]]
the good he instructed that they in their goodness through-lived

'To the sinful [...] he promised God's goodness [...]. The good, he instructed to keep their goodness.'
(coneot,LS_28_[Neot]:57.48-49)

I will first consider the potential topic categories to be included in my grammar model. I assume that a prominent discourse role of a phrase, as illustrated in the examples above, correlates with their syntactic position in the specifier of CP. In contrast, unmarked clause-initial material, such as scene-setting, speaker-orientation, discourse filler, or interjections and vocatives, are treated as adjunction to CP. Strictly speaking, one should thus investigate the information-structural status of initial constituents to judge their syntactic position. However, this is not a feasible approach for this study. Firstly, a manual evaluation of the discourse role of all clause-initial constituents would take far too much time and effort to be commensurate with the goal of this investigation. Secondly, categorizing the information-structural status of a phrase is highly subjective. Hence, I decided to approximate the topic status of a clause-initial constituent solely with the formal annotations provided by the corpora that I use as the basis for this study.

(i) Adverb and prepositional phrases can be placed in front of elements that I assume to be in the specifier of CP, such as nominal topics, *wh*-pronouns in direct questions or operator adverbs (see (42a)), and are therefore treated as adjuncts in such contexts. It thus becomes difficult to ascertain if a clause-initial adverb or prepositional phrase really occurs in the designated topic position, the specifier of CP, or if it should be parsed as a clausal modifier, an adjunct to CP. I therefore decided to ignore these categories in my model.
(ii) Other potential topic categories occur very infrequently in clause-initial position. This is true, for instance, for prepositional phrases that are unambiguously extracted from a lower domain, as in (66b), adjective and quantifier phrases, verb phrase topicalization, or extracted nominal modifiers. Those kinds of syntactic categories were therefore not considered either.
(iii) This basically left only argument DPs as potential topics. They are both frequent enough to be suitable for my purposes and fairly unambiguously placed in the specifier of CP when positioned clause-initially. Those sentence elements will therefore constitute the only categories to be included in my formal grammar model.

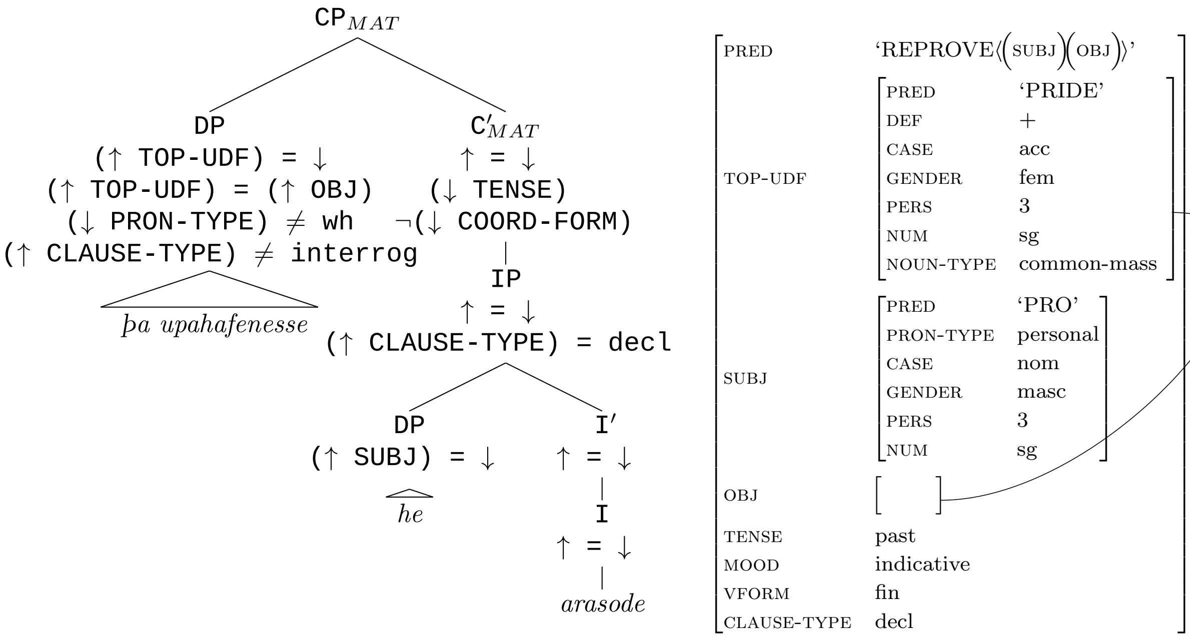
I will now show how my model analyzes topicalized argument DPs. Clause-initial object DPs are handled by the annotated phrase structure rule in (67). The rule will introduce an underspecified discourse function for a topic, which I label with the mnemonic TOP-UDF, in the specifier of CP. It must be unified with an arbitrarily deeply embedded direct object, OBJ, indirect object, OBJ$_{th}$, or extracted complement out of a local or long-distance oblique or adjunct PP resulting in preposition stranding. The rule can be applied only if the object is not headed by a *wh*-item and if the clause type attribute is not valued as interrogative.

(67) a. CP$_{MAT}$ → DP C′

| DP | C′ |
|---|---|
| (↑ TOP-UDF) = ↓ | ↑ = ↓ |
| {(↑ TOP-UDF) = (↑ {XCOMP\|COMP}$^*$ OBJ) | (↓ TENSE) |
| \| (↑ TOP-UDF) = (↑ {XCOMP\|COMP}$^*$ OBJ$_{th}$)} | ¬(↓ COORD-FORM) |
| \|(↑ TOP-UDF) = (↑ {XCOMP\|COMP}$^*$ {OBJ$_{th}$\|ADJUNCT} OBJ$_{th}$) ¬ (→ TENSE) } | |
| (↓ PRON-TYPE) ≠ wh | |
| (↑ CLAUSE-TYPE) ≠ interrog | |

The following example illustrates how the model parses topicalized objects.

(68) þa upahafenesse he arasode
the up-heaved-ness he reproved
'The pride, he reproved' (cocura,CP:4.39.20.210)




The example involves an object in front of a pronominal subject. Since my grammar model requires pronominal subjects to be elements in the specifier of IP, the object must have topicalized and can therefore only be parsed with the topicalization rule in (67). Specifically, the functional uncertainty constraint can only be realized as (↑ TOP-UDF) = (↑ OBJ) to arrive at a complete and coherent parse. The clause-initial underspecified discourse function thus gets unified with the object argument. The node C′$_{MAT}$ directly re-writes to IP, thereby signaling a declarative clause reading. Furthermore, the topicalized object is not headed by a *wh*-element. Therefore, all the constraints imposed by topicalization are met. The remainder of the sentence is analyzed with mechanisms explained earlier. The model produces one solution for the parse (technically two because *arasode* can be indicative or subjunctive).

While topicalization is impossible in interrogatives, it remains compatible with non-canonical declaratives. My model can therefore generate clauses with topicalization and simultaneously a verb that is independently licensed under C because of negative polarity or non-indicative mood. A few examples are given below.

(69) a. [$_{CP}$ [$_{DP}$ **nane are** ] [$_{C'}$ *ne dyde* [$_{P}$ he him ]]]
no honor not did he them
'No good deeds, he performed for them (coverhom,HomU_9_[ScraggVerc_4]:224.731)

b. [$_{CP}$ [$_{DP}$ **LX scillinga** ] [$_{C'}$ *gebete* [$_{IP}$ he ]]]
40 shillings better.subjunctive he
'{ May he / Let him / He should } make amends with 40 shillings' (colawine,LawIne:6.2.21)

c. [$_{CP}$ [$_{CP}$ **þæt flæsc þæt wildeor abiton**, ] [$_{C'}$ *ne ete* [$_{IP}$ ge ]]]
that flesh that wild-animals off-bite not eat.subjunctive/imperative? you
'The flesh that animals partake of, you should not eat' (cootest,Exod:22.31.3296)

Next, I will turn to the option for subject topicalization. It is not easy to determine if subjects can topicalize to begin with because the string resulting from topicalization, 'subject - verb,' would be identical to strings in I-initial main clauses with the subject in the specifier of IP, i.e., also 'subject - verb.' Nevertheless it is quite likely that parses are available, and in fact common, in which non-pronominal subjects are positioned in the C-domain. I will briefly mention four arguments in favor of this view.

(i) There is ample evidence for the existence of a discourse function in the clausal left periphery through fronted non-subjects. It seems intuitive for speakers to extend the functional licensing conditions on this position to all grammatical functions, including subjects.

(ii) Old English verbs have a very high probability of being placed under C if they are both negative and subjunctive. Hence, a subject preceding such a verb is probably positioned in the specifier of CP, as in (70).

(70) [$_{CP}$ [$_{DP}$ Ræpsas ] [$_{C'}$ **ne syn** [$_{IP}$ gesungene mid Alleluian ] ] ]
reply not be.subjunctive sung with hallelujah

'A response (in church service) should not be sung with a hallelujah'
(cobenrul,BenR:15.39.20.518)

(iii) Non-subject pronouns often occur in the left periphery of the clause in Old English. In the next section, I will argue that they often indicate the IP-boundary in this context. A subject preceding such a preposed pronoun is therefore likely to be positioned in the specifier of CP. This is illustrated in (71).

(71) [$_{CP}$ [$_{DP}$ min God ] [$_{IP}$ **me** asende to _ sona his engel ] ]
my God me sent to soon his angel

'My God sent his angel to me at once'
(coaelhom,ÆHom_22:326.3470)

(iv) Finally, there are a few attestations of topicalized subjects extracted from embedded clauses. Such subjects are easily parsed as elements in the specifier of CP but difficult to account for otherwise. Example (72a) could perhaps involve a quotative parenthetical. Example (72b) illustrates occasional violations of the *that*-trace effect (e.g., Perlmutter 1971) in Old English, licensed, perhaps, by the fronting of an adjunct inside the embedded clause. It thus seems to necessitate an analysis with subject topicalization. Below, I indicate subject topicalization parses with labeled bracketing and informally sketch the dependency between the subject topic and the subject of the embedded clause with an arrow.

(72) a. [$_{CP}$ [$_{DP}$ Sio winestre hand Godes ] [$_{IP}$ he cwæð [$_{CP}$ _ wære under his heafde ]]]
the left hand of-God he said were under his head
'God's left hand, he said was under his head' (cocura,CP:50.389.11.2635)

b. [$_{CP}$ [$_{DP}$ þæs regul ] [$_{IP}$ ic wille, [$_{CP}$ þæt gesinlice _ sie geræd on geferrædenne ]]]
this rule I want that diligently be read in company
'This rule, I want to be read out diligently in company' (cobenrul,BenR:66.127.9.1223)

Assuming then that subject topicalization is indeed a grammatical option in Old English, I will incorporate the rule for subject topicalization presented in (73) into my model. It is identical to the rule for topicalized objects except that it cannot be applied to pronominal elements. In this way, I preserve the diagnostic value of pronoun subjects for the specifier of IP.

(73) a. $\mathtt{CP}_{MAT} \rightarrow$

```
                    DP                              C′
            (↑ TOP-UDF) = ↓                       ↑ = ↓
{(↑ TOP-UDF) = (↑ {XCOMP|COMP}* SUBJ)         (↓ TENSE)
            (↓ PRON-TYPE) ≠ wh              ¬(↓ COORD-FORM)
        (↑ CLAUSE-TYPE) ≠ interrog
            (↓ PRED) ≠ PRO
```

A consequence of the admission of rule (73) into the grammar is that 'subject - verb' structures become ambiguous. They can now place the subject in its canonical position, the specifier of IP, or involve topicalization of the subject to the specifier of CP. This is illustrated in (74).

(74) a. **God** lufað us.
God loves us

'God loves us' (cocathom1,ÆCHom_I,_19:331.179.3761)

b. *Parse with subject in Spec,IP*
[$_{CP}$ [$_{IP}$ **God** [$_{I'}$ lufað us ]]]

c. *Parse with subject in Spec,CP*
[$_{CP}$ **God** [$_{IP}$ [$_{I'}$ lufað us ]]] ]

Next, I will discuss the diachronic development of topicalization in the history of English. The diachrony of Old English topicalization is frequently discussed not in isolation but with reference to the position of a nominal subject. Topics can either immediately precede a nominal subject, as in (75a), or be followed by the finite verb and a subsequent nominal subject, as in (75b), creating verb-second patterns (see also (19)). The parses indicated with labeled brackets make use of I-initial headedness typical of main clauses.

(75) a. Topic - Subject - Verb

[$_{CP}$ Đone cwide [$_{IP}$ Paulus [$_{I'}$ *geryhte* eft to biscepum ]]]
that saying Paulus applied again to bishops

'This saying, paul applied again to certain bishops' (cocura,CP:16.105.8.691)

b. Topic - Verb - Subject

[$_{CP}$ Đas getacnunge [$_{IP}$ *unwreah* [$_{VP}$ se ylca Iohannes mid þysum wordum ]]]
this betokening revealed the same Johannes with these words

'John revealed this signification with these words' (cocathom1,ÆCHom_I,_25:383.120.4889)

Unfortunately, the frequent conceptualization of topicalization in terms of verb-second or verb-third orders blends together several aspects of Early English syntax that are actually quite distinct - the nature of the clause-initial topic position, placement options for full subjects in a higher or lower position, and re-write rules for finite verbs under I or C. All of those aspects undergo substantial changes in Old and Middle English. If topicalization is measured in terms of the presence or absence of full subject - verb inversion, a relatively complicated trajectory emerges. In Old English, verb-second seems to be stable or perhaps slightly increasing (Haeberli 2002*b*: 251). In Middle English, the syntactic determinants of the variation between the two word order options become quite complex (Warner 2007), may increase due to a dialect split between southern and northern texts, where verbs may more freely occur before pronoun subjects, i.e., under C, in this context (Kroch and Taylor 1997), and may even be influenced by Anglo-Norman syntax (Haeberli 2010). In late Middle English, verb-second patterns sharply decline and virtually disappear by the seventeenth century (Fischer et al. 2000: 132-4), or rather stabilize at a very low frequency. I believe that the blurry picture of the history of verb-second is a reflection of the fact that it mixes together and describes the development of more than one diachronically variable pattern.

Hence, it might be more illuminating to investigate "topic-first syntax rather than verb-second" (Los 2012: 41). If topicalization is measured in terms of the presence or absence of a clause-initial topic, a clearer diachrony emerges. For example, Speyer (2010) shows that direct object topicalization out of all clauses with a direct object consistently declines from early Old to late Early Modern English. His graph is given in figure 3.3 below. One can clearly discern a decrease in nominal object topicalization from about 15% to about 1-5%.

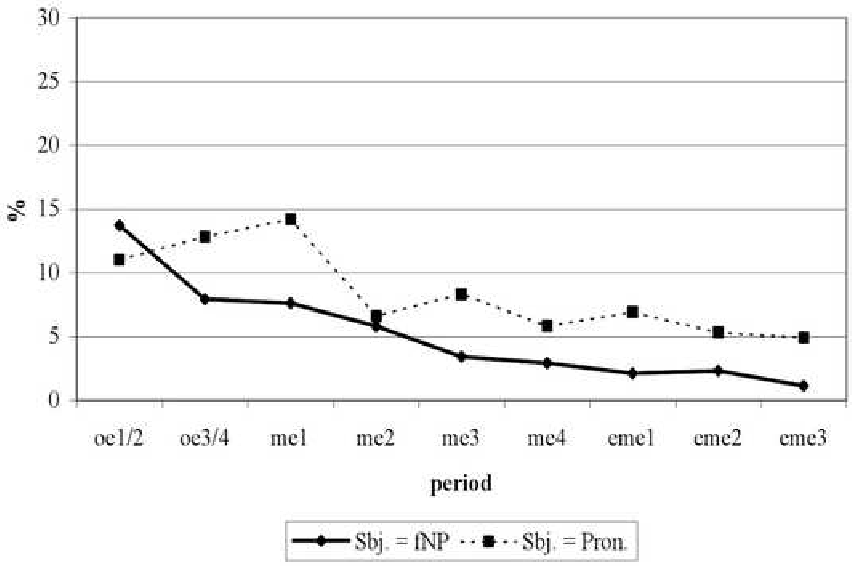



Figure 3.3: Rate of direct object topicalization in clauses with full (fNP) and pronominal subjects (Pron.) From: Speyer (2010: 52)

Thus, when one abstracts away from confounding factors caused by variable verb and subject positions, the employment of topicalization clearly seems to be losing ground. The general decline in argument topicalization can thus count as an additional, established change in Early English.

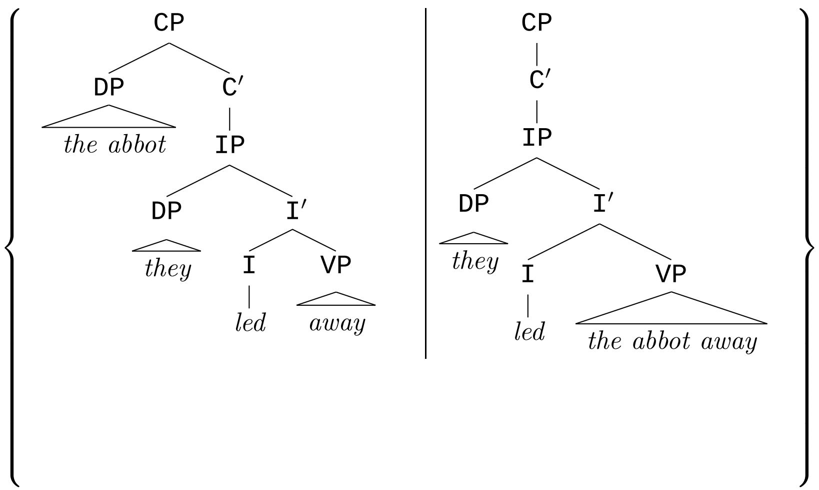
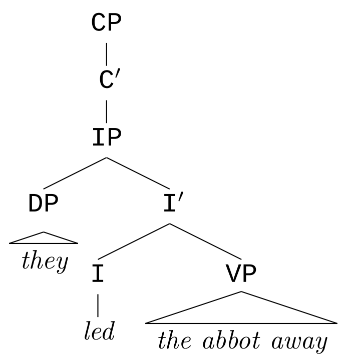
How should the decline in topicalization be conceptualized properly? There seems to be a fairly wide-spread consensus that it should not be viewed as an ordinary syntactic change involving competing variants and an eventual replacement of one phrase-structural alignment with another. This belief is bolstered by the fact that topicalization is still possible in Modern English. For instance, my grammar model for Modern English also includes a topicalization rule, formalized in rule (7c) of the previous chapter, which is remarkably similar to its Old English analogue in structure and syntactic licensing constraints. Hence, rather than a substitution of the syntactic rule itself, a change in (a) the information-structural or (b) phonological conditions under which it is used felicitously seems to have taken place.

(a) Information-structure. Most scholars prefer an explanation for the decline in topicalization in terms of a change in information-structural considerations that speakers have in mind. However, in my opinion, there is still no satisfactory understanding of the specific discursive licensing conditions of Old English topicalization or the exact information-structural differences between Old and Modern English topicalization. Dreschler (2015) sees the Old English topic position as a local anchoring position (233), argues that local anchoring declines with PPs, especially P+D clusters (*with that, for þy* etc.) and non-subject personal pronouns (301), and attempts to extend this analysis to object fronting as well (317). Similarly, Speyer (2010: 37-43) believes that the decline in Early English argument topics concerns mainly *e*[ntity] aboutness topics. His proxy construction for aboutness topics are fronted non-subject pronouns. Light (2012) contributes several related observations on the information structure of Old English fronted objects, especially bare demonstratives. Los (2009), formulating her generalizations in terms of the loss of verb-second, cites a number of information-structurally salient constructions that may have taken over the original function of Old English topicalization, such as clefts, passives or hanging topic left-dislocations. A special kind of aboutness-shift topic licensing null subjects may have been lost during Old English (Walkden 2014, Gelderen 2013).

(b) Prosody. Speyer (2010) proposes that the loss of inversion with full nominal subects after topicalized constituents (as in (75b)) would have led to unacceptable stress patterns between topic and full subject (??*Béans, Jóhn likes*). Topicalization was therefore avoided. In my opinion, this account is problematic. (i) There is no evidence for the proposed causal directionality between the declines in topicalization and verb-second. Rather than the loss of subject-verb inversion resulting in reduced topicalization, a general decline in topicalization could plausibly lead to fewer verb-second structures. Alternatively, a third factor could have caused both. (ii) Topicalization of full direct objects declines with both full and pronominal subjects. This seems to me to be unexpected under Speyer's account. (iii) Why does the prosodic well-formedness constraint to avoid two stresses / foci in a row outrank the applicability of a syntactic rule? In principle, there is nothing wrong with spondees (*Útáh, shút úp*). (iv) Heavy topics with an appropriate phonological structure should still frequently topicalize. For example, sentences like *Móckingbirds, Jóhn likes* should have persisted at original rates of topicalization. (v) It would seem to be extremely difficult to determine on the basis of historical text data if a topic, say a pronoun or demonstrative, was actually stressed. These problems led me to ignore the possible role of prosody on the decline in topicalization for my formalization.

I will follow the majority view on this issue and assume that it was a change in the information-structural conditions on topicalization that resulted in a reduced number of topicalized arguments overall. I will not attempt to clarify in any way what exactly these discursive changes might have been. Instead, I will simply assume that there is some specific discourse context X that regularly used to lead to the application of topicalization in Old English but no longer does so in Modern English. I can then model competition between topicalization and the absence of topicalization within the set of utterances that correspond to this relevant discourse context X. The figure below illustrates this approach. An arrow from "discourse context X" towards the variable syntactic structures indicates the restriction of syntactic competition to a particular information-structural alignment of the grammatical functions. It is not clear if there was ever a stage in the English language where the stipulated discourse context X obligatorily required topicalization and thus I leave out the conservative variant with a question mark. The requirement to condition different syntactic variants of a linguistic variable on an particular information-structural alignment stretches the model of competitive syntactic changes described in chapter 1 to its limits. Nevertheless, I will work with it as an approximation of the actual changes that took place during the Early English period in the remainder of the chapter.

**Change in Topicalization**

| | | | | |
|---|---|---|---|---|
| Conservative period: Speakers represent only the conservative variant $\alpha$ | | Transitional period: Speakers represent both the conservative variant $\alpha$ and the innovative variant $\beta$ | | Innovative period: Speakers represent only the innovative variant $\beta$ |
| $\alpha$ | $\rightarrow$ | $\{\ \alpha \mid \beta\ \}$ | $\rightarrow$ | $\beta$ |
| Conservative period: Unknown if speakers ever obligatorily required topicalization in a certain discourse context X | | Transitional period: Speakers allow both argument topicalization and the absence of topicalization in a certain discourse context X | | Innovative period: Speakers require the absence of topicalization in a certain discourse context X |
| | | discourse context X $\downarrow$ \| discourse context X $\downarrow$ | | discourse context X $\downarrow$ |
| ? | |  | $\rightarrow$ |  |

#### 3.3.1.7 Preposing of Pronouns

Pronouns in Old English frequently prepose. In particular, they can occur in two positions within the IP-layer. I will illustrate with subordinate clauses to exclude topicalization of the nominal subject as a parsing option.

Firstly, they can occur between a subject in the specifier of IP and the finite verb. Examples are presented in (76). The first example does not necessarily involve a pronoun in the I-domain. The reason is that it could be parsed with I-final phrase structure and extraposition of the post-verbal material so that the pronoun could remain in the VP. Most clauses with pronouns between subject and verb have this structure. However, constructions of this type are so common that it does not seem plausible to assume I-final phrase structure for all of them. Furthermore, the subsequent examples involve post-verbal non-postposing particles or stranded prepositions, demonstrating that at least some cases must be parsed with I-initial headedness, and hence the pronoun in the I-domain. The last example is taken from the Middle English verse text *Poema Morale*.

(76) *Complementizer - Subject - Pronoun - Verb*

a. ... [$_{\text{CP}}$ gyf [$_{\text{IP}}$ hwylc man **hyne** *begyrdeþ* mid þysse wyrte ]]
... if which man him clothes with this plant
'... if any man wears this plant' (coherbar,Lch_I_[Herb]:90.14.1471)

b. ... [$_{\text{CP}}$ swa þæt [$_{\text{IP}}$ þæt spere **him** *eode* [$_{\text{VP}}$ þurh ut ]]]
... so that that spear him went through out
'... so that the spear pierced through him' (coaelive,ÆLS[Ash_Wed]:50.2731)

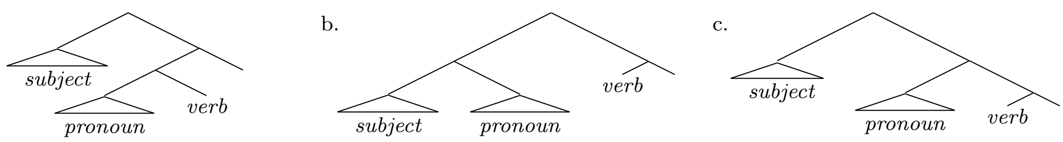
c. ... [$_{\text{CP}}$ oð [$_{\text{IP}}$ heora wif **him** *sendon* [$_{\text{VP}}$ ærendracan æfter ]]]
... until their wives them send messenger after
'... until their wives sent a messenger to them' (coorosiu,Or_1:10.29.7.564)

d. ... al [$_{\text{CP}}$ þat [$_{\text{IP}}$ þe loðe gost **hem** *tihte* [$_{\text{VP}}$ to ]] and taihte ]
... al that the loathsome ghost them enticed to and taught
'... all that the loathsome spirit enticed them to, and taught them.'
(PoemaMorale,229.228.276)

There is a substantial literature on the analysis of preposed pronouns of this kind. Abstracting away from specific proposals, there seem to be three options for the parse of the string 'subject - pronoun - verb,' which I will describe in terms of label-free, bare phrase structure. First, the pronoun and the verb could form a constituent to the exclusion of the subject and the rest of the clause, (78a). Such an analysis has been proposed, for instance, by van Kemande (1987: ch. 4). She assumes that pronouns can be clitics that adjoin to the finite verb. However, this structure would predict that the pronoun+verb cluster could also occur under C so that one would generate structures such as '*wh*-item - pronoun - verb - subject' for direct questions, contrary to fact. Secondly, one could let subject and pronoun form a constituent to the exclusion of the finite verb, (78b). A structure of this kind has been suggested, for example, by Pintzuk (1996), who adjoins pronouns as phrasal affixes to the left or the right to a topic in the specifier of IP, and more recently by van Kemade and Milicev (2011), who use the adverbs *þa, þonne* rather than I-initial finite verbs as a relevant demarcation category and allow both nominal subjects and non-subject pronouns to occur inside a topic Σ-phrase, which leads to the same hierarchy. This constituency, too, runs into certain empirical problems. I will mention just one. One would expect that the subject+pronoun cluster could be coordinated, which does not seem to be true. Coordination of a node containing pronoun and finite verb to the exclusion of the subject, on the other hand, is attested, as illustrated below. The subject+pronoun analysis does not make a position for the pronoun available.

(77) ... buton us Drihten Crist oþþe [$_{\text{I}'}$ [$_{\text{VP}}$ his leoht ] forgife ], oþþe [$_{?}$ **us** *læde* [$_{\text{VP}}$ onweg ]]
... unless us Lord Christ either his light give or us lead away
'... unless Lord Jesus Christ either gave us his light or led us away' (coaelive,ÆLS_[Maurice]:107.5743)

The third option would let the pronoun dominate the node containing the verb and the rest of the clause, forming a constituent to the exclusion of the subject (78c). This constituency avoids the previous problems and I will therefore use it in my formalization.

(78) a.


The rule in (79) will adjoin pronouns to I′ thus creating the desired constituent structure. The pronoun is mapped onto an underspecified discourse function used specifically for pronouns, abbreviated PRO-UDF, which can be unified with direct objects, indirect objects or prepositional complements. The restriction to pronouns is achieved by constraining the DP's PRED feature.

(79) I′ → DP I′
(↑ PRO-UDF) = ↓ ↑ = ↓
{ (↑ PRO-UDF) = (↑ OBJ)
|(↑ PRO-UDF) = (↑ OBJ$_{th}$)
|(↑ PRO-UDF) = (↑ {OBJ$_{th}$|ADJUNCT} OBJ$_{th}$)
¬(→ TENSE) }
(↓ PRED) =$_c$ PRO

The second option for pronominal preposing places a pronoun at the CP-IP boundary, so that it is sandwiched between C and a subject in the specifier of IP. This is illustrated in (80). The first example involves a full subject, which could also be parsed in the lower subject position with I-final headedness and extraposition of the post-verbal material. In this case, the pronoun might be positioned lower in the structure as well. The majority of examples are of this type. However, they are numerous enough to make an I-final analysis for all of them improbable. Furthermore, the subsequent three examples all include post-verbal stranded prepositions, which cannot postpose, thus requiring I-initial headedness so that the pronoun must have preposed in these cases. The second example illustrates a fairly common structure in which a preposed pronouns immediately precedes the indefinite subject pronoun *man*. In contrast, preposed pronouns do not precede personal pronoun subjects in such alignments. The post-pronominal subject can also be non-pronominal, as in the third example. The final example is taken from the Middle English *Ormulum*.

(80) *Complementizer - Pronoun - Subject - Verb*

a. ... [$_{\text{CP}}$ oþ þæt [$_{\text{IP}}$ **hiene** an swan *ofstang* æt Pryfetes flodan ]]
... until that him a swain stabbed at Privett's stream
'... until a swineherd stabbed him to death at Privett's stream'
(cochronA-CC,ChronA_[Plummer]:755.1.509)

b. ... [$_{\text{CP}}$ þeah ðe [$_{\text{IP}}$ **þe** man *bere* [$_{\text{VP}}$ mete toforan ]]]
... though that you one carry food before
'... though one might bring food before you' (cocathom2,ÆCHom_II,_29:234.117.5202)

c. Eft gif heo wylle þæt [$_{\text{CP}}$ ðæt [$_{\text{IP}}$ **hyre** blodryne *cyme* [$_{\text{VP}}$ to ]]] ...
again if she wants that [namely] that her blood-flux came to ...
'Again, if she desires that a bloody flux should come to her ...'
(coquadru,Med_1.1_[de_Vriend]:2.3.71)

d. ... Swillc haliȝ bisne [...] [$_{\text{CP}}$ Alls [$_{\text{IP}}$ **himm** hiss herrte *bereþþ* [$_{\text{VP}}$ to ]]]
... such holy example as him his heart carries to
'... such a holy example [...] to which his heart carries him' (CMORM,I,193.1582)

Following the same logic regarding the constituency of 'subject - pronoun - verb' structures, I will assume that preposed pronouns can also adjoin to IP. The relevant rule for this parse is presented in (81) below. It is identical to the rule in (79), but adds a conditional constraint to rule out pronominal personal subjects co-occuring with the preposed pronoun

(81) IP → DP IP
(↑ PRO-UDF) = ↓ ↑ = ↓
{ (↑ PRO-UDF) = (↑ OBJ)
|(↑ PRO-UDF) = (↑ OBJ$_{th}$)
|(↑ PRO-UDF) = (↑ OBJ$_{th}$)
|(↑ PRO-UDF) = (↑ {OBJ$_{th}$|ADJUNCT} OBJ$_{th}$)
¬(→ TENSE) }
(↓ PRED) =$_c$ PRO
¬(↑ SUBJECT PRON-TYPE) = PERSONAL

My model can now handle a wide range of structures involving pronoun preposing to the IP-domain. Unfortunately, the wider coverage comes at the expense of substantially increased structural ambiguity. I shall give two examples.

First, the pattern 'subject pronoun - pronoun - verb' has become ambiguous. An example was given in (27a), which I repeat in (82a) for convenience. Up to this point, this sentence allowed only an `I`-final structure with the pronoun inside the `VP`, as explicitly shown in (29). This parse is repeated in (82b). The introduction of rule (79) now permits a second parse, namely the preposing of the pronoun to `I'`. This second parse is sketched in (82c).

(82) a. he **hig** þreade
he them rebuked
'He rebuked them' (cowsgosp,Lk_[WSCp]:9.55.4439)

b. *`I`-final parse with pronoun inside `VP`*
[$_{\texttt{IP}}$ he [$_{\texttt{I'}}$ [$_{\texttt{VP}}$ **hig** ] þreade ]]

c. *Parse with preposing of the pronoun to `I'`*
[$_{\texttt{IP}}$ he [$_{\texttt{I'}}$ **hig** [$_{\texttt{I'}}$ þreade ]]]

The second example shows a clause-initial non-subject pronoun in a verb-second clause. An example is presented in (83a). Such structures can now be parsed in three different ways. They could involve object topicalization to the specifier of `CP`, (83b), pronominal preposing to `IP`, (83c), or preposing of the pronoun to `I'`, (83d).

(83) a. **Him** cwæð Nichodemus to
him spoke Nichodemus to
'Nichodemus spoke to him' (coaelhom,ÆHom_13:10.1886)

b. *Parse with object topicalization*
[$_{\texttt{CP}}$ **Him** [$_{\texttt{IP}}$ [$_{\texttt{I'}}$ cwæð [$_{\texttt{VP}}$ Nichodemus to ] ]]]

c. *Parse with pronominal preposing to `IP`*
[$_{\texttt{CP}}$ [$_{\texttt{IP}}$ **Him** [$_{\texttt{IP}}$ [$_{\texttt{I'}}$ cwæð [$_{\texttt{VP}}$ Nichodemus to ] ]]]]

d. *Parse with pronominal preposing to `I'`*
[$_{\texttt{CP}}$ [$_{\texttt{IP}}$ [$_{\texttt{I'}}$ **Him** [$_{\texttt{I'}}$ cwæð [$_{\texttt{VP}}$ Nichodemus to ] ]]]]

This ambiguity has consequences for the measurement of topicalization with pronominal objects. Speyer (2010: ch. 2.2), for instance, uses the rate of fronted object pronouns as a proxy for his category of *e*[ntity] topics in order to back up his claim that it is mainly a decline in those *e*-topics that led to a decline in Early English topicalization overall. What one really measures by considering the sharp decline in fronted object pronouns (in Speyer's data: 11.0% in OE1/2, 6.6%in OE3/4, 3.9%in ME1) may not be a decline in topicalization, but a decline in pronominal preposing. In fact, I would submit that pronominal objects may not have topicalized in Old English any more frequently than they do in Modern English, i.e., at a negligibly low rate, perhaps owing to their low deictic value. Rather, the majority of Old English fronted pronouns could plausibly be generated by one of the preposing rules. However, since many scholars assume that pronominal topicalization is a genuine grammatical option in Old English, I will not restrict my model in any way to reduce its potential for structural ambiguity, and leave the development of arguments against pronoun topicalization in Old English for future research.

#### 3.3.1.8 Dual Category Membership of Conjunctions

I will now turn to the formalization of coordination in Old English. At the heart of my proposal is the idea that Old English conjunctions can belong to two distinct syntactic categories. Consequently, different phrase structure rules will be responsible for their insertion into the constituent structure. One cannot unequivocally decide which category a given conjunction should belong to on the basis of its formal appearance or by its function. Rather, their multiple word class membership is revealed indirectly through observations of statistically significant differences in word order distributions. I will first formalize these two conjunction types. Subsequently, I will attempt to present some independent arguments for their multi-class membership.

The two syntactic categories of Old English conjunctions are the following. (i) They may belong to the proper syntactic category of conjunctions as in Modern English, labeled `CONJ`. I will call these conjunctions *logical connectors*. Logical connectors bear a coordination form feature, `COORD-FORM`, which does not distribute over a set of conjuncts. Lexicon entries for the three Old English logical connectors *and*, *ac* and *ne* are shown in (84).

(84) a. *and* CONJ (↑ COORD-FORM)= and
b. *ac* CONJ (↑ COORD-FORM)= but
c. *ne* CONJ (↑ COORD-FORM)= nor

(ii) Alternatively, Old English conjunctions can be of the same type as complementizers, i.e., C. I therefore refer to those items as *C-head conjunctions*. Such C-head conjunctions are modeled as synonymous to logical connectors; they also introduce a COORD-FORM feature. Lexicon entries for the three Old English conjunctions *and*, *ac* and *ne* as C-head conjunctions are presented in (85).

(85) a. *and* $\text{C}_{MAT}$ (↑ COORD-FORM)= and
b. *ac* $\text{C}_{MAT}$ (↑ COORD-FORM)= but
c. *ne* $\text{C}_{MAT}$ (↑ COORD-FORM)= nor

Logical connectors are targeted by a syntactic rule for sentential coordination. I will adopt a version of this rule that is in line with the most widely adopted treatment of coordination in the LFG literature (e.g., Dalrymple 2001: ch. 13, Sadler 2005). The rule builds a coordinate structure consisting of one or more conjunct CPs before the conjunction, indicated by a Kleene plus sign, '+,' the logical connector itself, and exactly one final conjunct CP. Every conjunct clause is interpreted as a member of a set that is contained in the f-structure denoted by the mother CP-node. Since the COORD-FORM-feature on conjunctions is non-distributive, it will be associated directly with the f-structure containing the set of conjuncts, not with the individual members of the conjunct set. The logical connector rule is specified in (86).

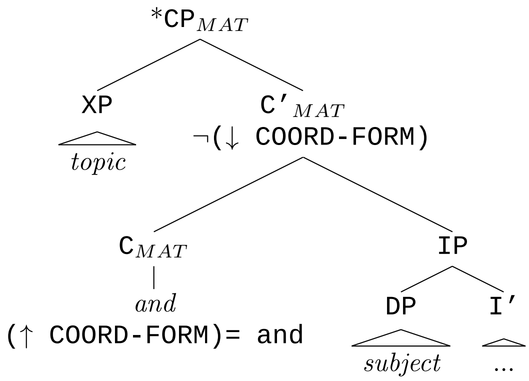
(86) $\text{CP}_{MAT} \rightarrow \underset{\downarrow \in \uparrow}{\text{CP}_{MAT}{}^{+}} \quad \underset{\uparrow = \downarrow}{\text{CONJ}} \quad \underset{\downarrow \in \uparrow}{\text{CP}_{MAT}}$

C-head conjunctions, on the other hand, are directly inserted under C by virtue of their syntactic category. The phrase structure rule responsible for the introduction of C-head conjunctions permits the introduction of a C-head in declarative matrix clauses if the C-head contains a COORD-FORM-feature. The C-head conjunction rule is shown in (87).

(87) $\text{C'}_{MAT} \rightarrow \begin{array}{c}\text{C}_{MAT}\\ \uparrow = \downarrow\\ (\downarrow \text{COORD-FORM})\\ (\uparrow \text{CLAUSE-TYPE}) = \text{decl}\end{array} \quad \begin{array}{c}\text{IP}\\ \uparrow = \downarrow\end{array}$

I constructed my toy grammar in such a way as to rule out parses with a topic in the specifier of CP and, simultaneously, a conjunction under C. The topicalization rules developed in section 3.3.1.6 involve a negative existential constraint, ¬(↓ COORD-FORM) on C', which conflicts with the feature specification of C-head conjunctions. Hence, C-head conjunctions will be incompatible with the topicalization rules and, as desired, main clause structures of the form 'topic - conjunction - subject' cannot be generated. Hence, C-head conjunction will disable the clausal left periphery for topicalization. This is sketched in (88).

(88)




It might be objected that this constraint makes the model too restrictive. Perhaps topicalization and C-head conjunctions can in fact be applied at the same time. While I could not find examples of conjunctions preceded by elements that would typically topicalize, such as direct or indirect objects, the YCOE does annotate a few sentences with clause-initial, left-dislocated noun phrases followed by a potential C-head conjunction. If left-dislocated noun phrases are placed in the topic position, they would indicate parses conforming to the banned structure in (88). For example, (89) includes a nominative left-dislocation, *John the evangelist*, followed by a conjunction, *and*, as well as a the subject, *Christ*. I indicate with labeled bracketing a parse that would contradict the ban on topicalization under the presence of a C-head conjunction.

(89) [$_{CP}$ [$_{DP}$ Iohannes se godspellere, þe Gode wæs gecweme, ]
John the evangelist, who to-God was pleasant
[$_{C'}$ **and** [$_{IP}$ Crist hine lufode for þære clænnysse ]]]
and Christ him loved for the cleanness
'And as for St. John, who was pleasant to God, Christ loved him for his chastity'
(colsigef,ÆLet_5_[Sigefyrth]:26.11)

It may also be noteworthy that parallel constructions can be found in subordinate clauses, where a topic is placed in front of a complementizer, as in (90). Constituent fronting to the C-domain that includes an overt C-head could thus be argued to be a commonality between CCs and subordinate clauses.

(90) Wite se abbod þeah,
know.subjunctive the abbot however
[$_{CP}$ [$_{DP}$ eal þæt he do ] [$_{C'}$ **þæt** [$_{IP}$ he hit do mid Godes ege ] ] ]
all that he does that he it does with God's awe
'The abbot should know, however, that he should do all that he does in the fear of God'
(cobenrul,BenR:3.16.5.231)

However, there are several reasons that make the structure indicated in (89) relatively unlikely. (i) It is plausible that left-dislocated elements are not topics in the specifier of CP. Rather, they occur in a structurally higher position. This is shown by the existence of sentences that place a left-dislocated noun phrase in front of a topic, both with resumptive subjects, (91a), or resumptive objects, (91b). Left-dislocations could thus potentially co-occur with a C-head conjunction without violating the constraint sketched in (88).

(91) a. [ Milred bisceop ]$_i$ [$_{CP}$ þare halegan rode tacen [$_{IP}$ he$_i$ heron gefæstnode ]]
Milred bishop the holy rood's token he hereon affixed
'As for bishop Milred, he affixed the sign of the Holy Rood hereon' (codocu3,Ch_98_[Rob_1]:19.32)

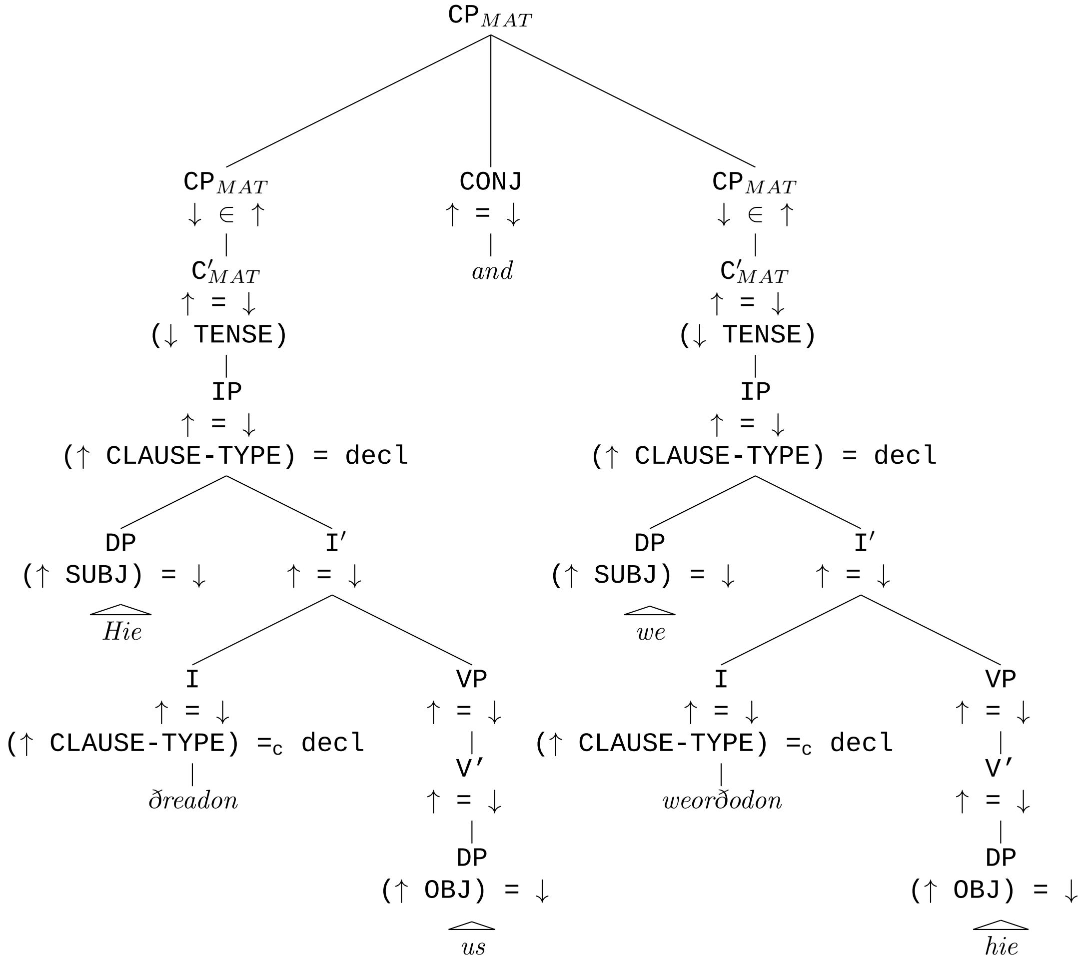
b. [ þam þe Dryhten mycel to forlæteð ]$_i$ [$_{CP}$ myceles [$_{IP}$ he hine$_i$ eac eft manað ]]
that-one who Lord much to grants much he him also again demands
'As for those to whom the Lord grants much, he also demands much from them in return'
(coverhom,HomS_40.3_[ScraggVerc_10]:204.1538)

(ii) More importantly, the sentences that seem to allow a left-dislocated topic to co-occur with a C-head conjunction may not have been annotated in the most expedient way in the YCOE. Reasonable, alternative parses remain available for the potential counterexamples. For instance, (89) could be analyzed with conjunction of the relative clause so that the entire string would be one left-dislocated noun phrase resumed by the pronominal subject of the following sentence.[2] (iii) Similarly, even if the annotation represents the most appropriate analysis of configurations as in (89), these structures could also result from speech disfluencies, manuscript copying errors or Latin interference effects. (iv) Finally, structures like (89), for which I found only 5 relevant instances, are considerably rarer than structures like (90), for which there may be about 100 instances in the surviving Old English texts. This difference would be surprising if simultaneous application of topicalization and C-head conjunctions were a genuine grammatical option in Old English. Thus, it is plausible to maintain the assumption that structures with a topic in the specifier of CP and a conjunction under C are in fact ungrammatical.

Old English constructions of the form 'clause - conjunction - clause' will now usually be ambiguous. They can be analyzed as a coordinate structure with a logical connector, as in (92), or as two independent clauses where the second occurs with a C-head conjunction, as shown in (93). This structural ambiguity is, in essence, the core feature of the theory of Old English coordination that I would like to propose in this chapter. As I will show, it can account for the distributional difference between MCs and CCs presented in the introduction in a straightforward way and predicts various additional synchronic and diachronic differences between the two clause types.

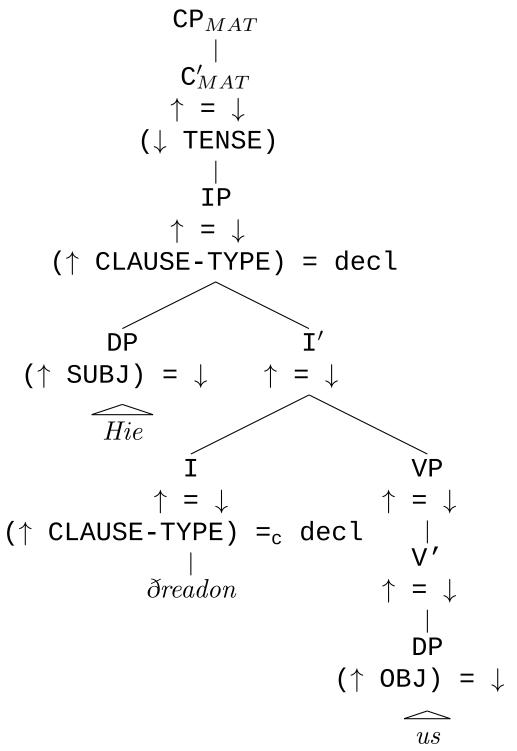
[2] I am indebted to Cynthia Allen for pointing out this alternative parse.

(92) Hie ðreadon us & we weorðodon hie
they rebuked us and we honored them
'They rebuked us and we honored them'
(cocura,CP:36.255.7.1666-1667)



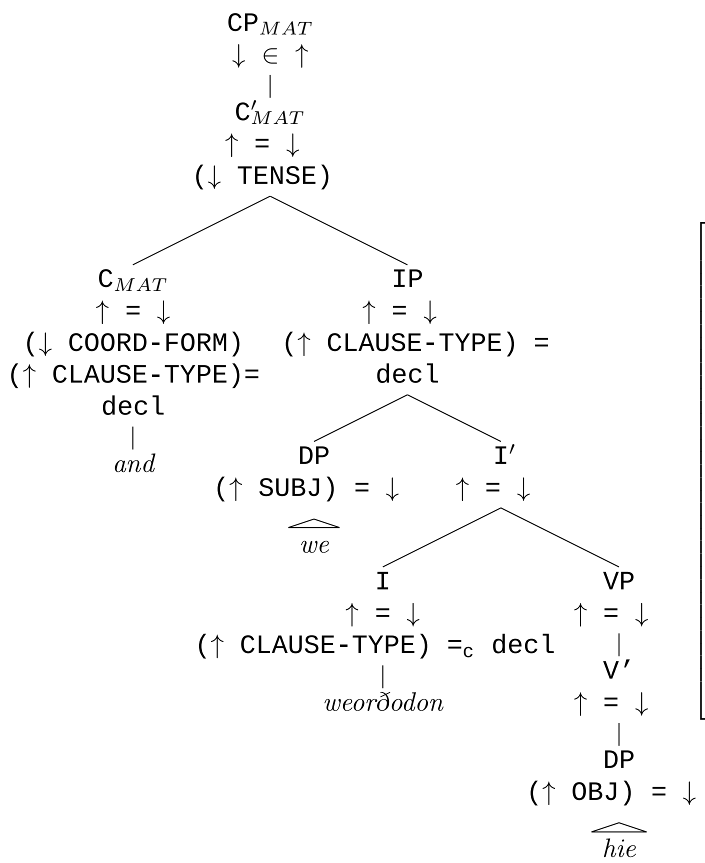

$$
\left[\begin{array}{ll}
\left\{\begin{array}{l}
\left[\begin{array}{ll}
\textsc{pred} & \text{`REBUKE}\langle(\textsc{subj})(\textsc{obj})\rangle\text{'} \\
\textsc{subj} & \left[\begin{array}{ll}\textsc{pred} & \text{`PRO'} \\ \textsc{pron-type} & \text{personal} \\ \textsc{case} & \text{nom} \\ \textsc{pers} & 3 \\ \textsc{num} & \text{pl}\end{array}\right] \\
\textsc{obj} & \left[\begin{array}{ll}\textsc{pred} & \text{`PRO'} \\ \textsc{pron-type} & \text{personal} \\ \textsc{case} & \text{acc} \\ \textsc{pers} & 1 \\ \textsc{num} & \text{pl}\end{array}\right] \\
\textsc{tense} & \text{past} \\
\textsc{mood} & \text{indicative} \\
\textsc{vform} & \text{fin} \\
\textsc{clause-type} & \text{decl}
\end{array}\right] \\
\left[\begin{array}{ll}
\textsc{pred} & \text{`HONOR}\langle(\textsc{subj})(\textsc{obj})\rangle\text{'} \\
\textsc{subj} & \left[\begin{array}{ll}\textsc{pred} & \text{`PRO'} \\ \textsc{pron-type} & \text{personal} \\ \textsc{case} & \text{nom} \\ \textsc{pers} & 1 \\ \textsc{num} & \text{pl}\end{array}\right] \\
\textsc{obj} & \left[\begin{array}{ll}\textsc{pred} & \text{`PRO'} \\ \textsc{pron-type} & \text{personal} \\ \textsc{case} & \text{acc} \\ \textsc{pers} & 3 \\ \textsc{num} & \text{pl}\end{array}\right] \\
\textsc{tense} & \text{past} \\
\textsc{mood} & \text{indicative} \\
\textsc{vform} & \text{fin} \\
\textsc{clause-type} & \text{decl}
\end{array}\right]
\end{array}\right\} & \\
\textsc{coord-form} & \text{and}
\end{array}\right]
$$

(93) a. Hie ðreadon us
they rebuked us
'They rebuked us' (cocura,CP:36.255.7.1666)




$$\begin{bmatrix} \textsc{pred} & \text{`REBUKE}\langle(\textsc{subj})(\textsc{obj})\rangle\text{'} \\ \textsc{subj} & \begin{bmatrix} \textsc{pred} & \text{`PRO'} \\ \textsc{pron-type} & \text{personal} \\ \textsc{case} & \text{nom} \\ \textsc{pers} & 3 \\ \textsc{num} & \text{pl} \end{bmatrix} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{`PRO'} \\ \textsc{pron-type} & \text{personal} \\ \textsc{case} & \text{acc} \\ \textsc{pers} & 1 \\ \textsc{num} & \text{pl} \end{bmatrix} \\ \textsc{tense} & \text{past} \\ \textsc{mood} & \text{indicative} \\ \textsc{vform} & \text{fin} \\ \textsc{clause-type} & \text{decl} \end{bmatrix}$$

b. & we weorðodon hie
and we honored them
'And we honored them' (cocura,CP:36.255.7.1667)




$$\begin{bmatrix} \textsc{pred} & \text{`HONOR}\langle(\textsc{subj})(\textsc{obj})\rangle\text{'} \\ \textsc{subj} & \begin{bmatrix} \textsc{pred} & \text{`PRO'} \\ \textsc{pron-type} & \text{personal} \\ \textsc{case} & \text{nom} \\ \textsc{pers} & 1 \\ \textsc{num} & \text{pl} \end{bmatrix} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{`PRO'} \\ \textsc{pron-type} & \text{personal} \\ \textsc{case} & \text{acc} \\ \textsc{pers} & 3 \\ \textsc{num} & \text{pl} \end{bmatrix} \\ \textsc{tense} & \text{past} \\ \textsc{mood} & \text{indicative} \\ \textsc{vform} & \text{fin} \\ \textsc{clause-type} & \text{decl} \\ \textsc{coord-form} & \text{and} \end{bmatrix}$$

I will now present several arguments in favor of the view that Old English conjunctions are ambiguous between logical connectors and C-heads. Unfortunately, these arguments are not particularly scientific as they cannot currently be tested in any meaningful way. Nevertheless, I decided to include them here because they are interesting in their own right and because they may contribute to a better understanding of my conceptualization of C-head conjunctions. My arguments concern (i) the etymology of Old English conjunctions, (ii) potential semantic differences between logical connectors and C-heads, and (iii) a particular syntactic construction that may indicate parses with a conjunction under `C`. My discussion will focus on the word *and* as it is by far the most common Old English conjunction.

**Etymology**

I will first discuss the etymology and a possible path for the development of the English word *and*.
(i) **Proto-Indo-European** The ancestor of *and* is usually given as Proto-Indo-European $^*$*$h_2$entí*, the locative singular of 'forehead, front', resulting in the meaning 'in front of, facing against' (e.g., Ringe 2006: 71). This yields, for instance, Latin *ante* 'in front of' as in the well-known Latin proverb *Hannibal* **ante** *portas* 'Hannibal is in front of the gates' or Ancient Greek ἀντί 'opposing,' the source for the Modern English neoclassical prefix *anti-* as in *antibiotics*. Hence, the original category of the word was probably a preposition.
(ii) **Proto-Germanic** In Proto-Germanic, the word became > $^*$*antí* after the loss of the *a*-coloring laryngeal *$h_2$*, > $^*$*anθí* by Grimm's Law, > $^*$*anðí* by Verner's Law, and eventually > $^*$*ánði* after the stress shift to the first syllable. The word also seems to have substantially changed its meaning as a preposition. The Gothic reflex *and* shows a number of polysemous, usually directional path, interpretations, such as <u>*and*</u> *allos aikklesjons* 'through all the churches' (Codex Ambrosianus A, Corinthians II 8:18), <u>*and*</u> *alla jaina airþa* 'into all that land' (Codex Argenteus, Matthew 9:26). It is not too difficult to imagine how such a semantic change might have come about. It can be understood in terms of the same cognitive processes by which the meaning of modern prepositions are believed to grammaticalize. For instance, the English adverse preposition *against* 'facing, opposite' as in *lean against a wall* can gather directional path semantics by focusing the frame on the directional path, as in *rise* **against** *the sky* (e.g., Rhee 2002). Such a development could have become generalized in Gothic.
The original meaning of opposition was preserved in Proto-Germanic as a versatile derivational prefix, *and-* (cf. German *ent-*). The prefix is involved in the formation of words such as Gothic *and-bindan* 'opposite-bind' > 'release, unfasten', Old English *and-swerian*, 'in return-swear' > 'answer,' and many more.
(iii) **North/West-Germanic**. In later stages of Germanic, the prepositional use of the word was likely lost. Instead, it now occurred without a nominal complement, i.e., as an adverb. The meaning of the adverb can be reconstructed as contrastive 'yet, nevertheless, still', additive 'in addition, once again, also' or temporal 'then, afterwards, and then.' The specific mechanism for the required category and semantic changes seem to be rather badly understood. One can perhaps hypothesize that the original opposing meaning was somehow reanalyzed on the basis of the adverse prefix. For example, if the original meaning of *answer* was construed as 'say in return,' there might be a path to the reinterpretation as 'say nevertheless,' 'say once again' or 'say afterwards.' Alternatively, there may have been influence from some irretrievable external source, such as a substrate language. Evidence for the adverbial stage of *and* comes mainly from North-Germanic, e.g., Old Norse *enn* 'still, yet' (Zoëga 1910), as in *ok varð* **enn** *sem fyrr* 'and [it] happened yet/just/again as before' (Prologue of Snorra Edda). The adverbial use of the word does not seem to be well-attested in West-Germanic. However, I would like to offer the following speculative argument in support of its existence: Old High German texts sometimes show apparent verb-first word order in conjunct declarative matrix clauses after *and*. This is unexpected since verb-first patterns are otherwise very rare in (late) Old High German texts (Axel 2007: 170). The word order can be explained by assuming that *and* actually functions as an adverb in these cases, meaning 'and then,' is placed in the specifier of `CP`, and thus actually counts as a normal verb-second clause. An example is shown below.

(94) *Old High German*

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| so | bizzet | siun | innan, | unzin | er | stirbit, | [$_{\text{CP}}$ [$_{\text{ADVP}}$ unde ] | [$_{\text{C}'}$ uerit | [$_{\text{IP}}$ siu | gesunt | uz.]]] |
| so | bites | she-him | inside, | until | he | dies, | and (then) | goes | she | sound | out |

'She [=the hydra] bites him [while she is] inside [of him], until he dies, and then she comes out sound.'
(Old High German Physiologus, Hydra)

(iv) **Anglo-Frisian**. It is my contention that the word *and* underwent an additional category change during its development towards Old English. Specifically, I would like to propose that the word followed the commonly assumed grammaticalization path from a specifier to a head (Roberts and Roussou 2003, Gelderen 2004). For example, the Greek *wh*-adverb ποῦ 'where, how' developed into the complementizer πού 'that.' Old English adverbial *and* may have developed into a complementizer in the same way.

(v) **Old English** Finally, it is easy to see that an item with contrastive, additive or temporal meaning could develop into a logical connector. Thus, a structure like [*X. Yet Y*] can become [*X but Y*]. Similarly, [*X. Also Y*] or [*X. Then Y*] can be reinterpreted as [*X and Y*]. Icelandic has the contrastive conjunction *en* 'but, and.' All West-Germanic languages continue *and* as an additive logical connector, English *and*, German *und*, Dutch *en* etc.
In summary, the English word *and* originated as a preposition P > became an adverb ADV > subsequently potentially a C-head C > and eventually a normal conjunction CONJ. Old English may record variation between the second to last and last stages in this development.

One can rationalize the complementizer status of other conjunctions with similar etymological arguments. For instance, *ac* 'but' cannot be traced back to Indo-European with certainty, but it likely originates as a Proto-Germanic weakening of an adverb meaning 'also', as shown by comparisons between Old English *ac* vs. *eac*, Gothic *ak* vs. *auk*, Old High German *oh* vs. *ouh* etc. (see Carr 1933 and references cited therein). Therefore, this word, too, seems to have undergone a grammaticalization process with characteristic phonological reduction from an adverb towards a more functional syntactic category. That this category change might have resulted in a complementizer can be advocated with Gothic data. Gothic has a frequent complementizer, *ei* 'that,' used as a relative particle, a subordinator for consecutive adverbial clauses, or, less commonly, a complementizer to introduce complement clauses, as in (95a). The Gothic word *ak* 'but' usually appears at the beginning of a clause, as in (95b). Now, Gothic also shows fusion of clause-initial *ak* with the complementizer *ei*. One could interpret this as grammaticalization of an adverb-like element in the specifier of CP towards a head in C, (95c). This process would be analogous to my above proposal for the Anglo-Frisian development of the word *and*.

(95) *Gothic*

a. þat ain wait [$_{\text{CP}}$ **ei** blinds was, iþ nu saihva ]
that one I-know that blind I-was, yet now I-see

'One thing I know, namely that I was blind, yet now I see' (Codex Argenteus, John 9:25)

b. **ak** lausei uns af þamma ubilin
but loose us of that evil

'But deliver us from evil' (Codex Argenteus, Matthew 6:13)

c. [$_{\text{CP}}$ [$_{\text{C}'}$ **akei** [$_{\text{IP}}$ ni ainshun uslagida ana ina handuns ]]]
but-that not anybody out-laid on him hands

'But nobody laid hands on him' (Codex Argenteus, John 7:44)

The Old English negative element *ne* could have undergone a similar development. Alternatively, the complementizer category of *and* could have been extended to other conjunctions by analogy.

In short, there are potential grammaticalization pathways for the ancestors of Old English conjunctions towards complementizer items. Again, I would like to stress that the mere possibility for the proposed development does not demonstrate its plausibility. My account is merely meant to provide a conceptual framework in which to situate the historical origin of the special status of Old English conjunct clauses.

#### Semantic Differences

Next, one might speculate that there should be meaning differences between logical connectors and C-head conjunctions. For example, the C-head conjunction *and* might preserve prominently the hypothesized adverbial sense of 'then, afterwards, also' whereas logical connector *and* could bear a more truth-conditional, logical, sense, 'the whole proposition is true if the individual propositions are true.' However, I was unable to objectively classify the meaning of *and* into more additive, temporal vs. more logical, connective senses. Both readings always seem to be available. Semantic distinctions seem to be even harder to conjecture for other Old English conjunctions.

Nevertheless, I will present some anecdotal evidence that may illustrate my speculative idea about potential meaning differences between C-heads and logical connectors. Specifically, I will show two passages from Old English texts that may exemplify semantic differences of the hypothesized kind.

The first example is a narrative passage from the Old English *Martyrology.* The series of introductory *and*s in (96b) - (96d) seem to me to be translatable quite naturally in a temporal sense as 'and then, afterwards, next.' Furthermore, they are compatible with syntactic parses under C - they are placed immediately before the subject and even show verb-final word order in (96c) and (96d). Thus, these instances of *and* may instantiate conservative C-head conjunctions.

(96) a. þa hi Stefanus to þære stæninge læddon þa mihte he locian on heofnas,
when they Stephen to the stoning led then could he look in heavens
'When they led Stephen to his stoning, then he could see into heaven'
(Mart_1_[Herzfeld-Kotzor]:De26,A.8.72)

b. [$_{CP}$ **ond** he geseah þone hælend silfne standan on his godþrimme. ]
and he saw the savior self stand in his god-majesty
'And he saw the Savior himself standing in his divine majesty'
(Mart_1_[Herzfeld-Kotzor]:De26,A.8.73)

c. [$_{CP}$ **Ond** he hit þam Iudeum sæde, ]
and he it the Jews said
'And he said it to the Jews '
(Mart_1_[Herzfeld-Kotzor]:De26,A.11.74)

d. [$_{CP}$ **ond** hi him micle þe reðran on wæron ]
and they him much the more-cruel on were
'And they become even more cruel towards him'
(Mart_1_[Herzfeld-Kotzor]:De26,A.11.75)

The second example comes from the *Anglo-Saxon Chronicle* entry for the year 1011. The sequence of sentences in (97b) - (97d) introduced by *and* can be interpreted, I think, as a mere list of events, and not necessarily as the temporal sequence of these events. Moreover, they may be better parsed as coordination structures - they never occur with verb-final orders and even allow for topicalization in (97c). Hence, these instances of *and* may realize innovative logical connectors.

(97) a. hi ymbsæton Cantwareburuh [...]
they around-set Canterbury [...]
'They besieged Canterbury. [...]'
(ChronC_[Rositzke]:1011.15.1518)

b. **And** [$_{CP}$ hi þær ða genaman þone arcebisceop Ælfeah and Ælfweard cynges gerefan and Leofrune abbatissam and Godwine bisceop, ]
And they there then took the archbishop Ælfeah and Ælfweard king's reeve and Leofrune abbess and Godwin bishop
'And there they then took the archbishop Æleah and Ælfwear, the king's reeve, and abbess Leofrun, and bishop Godwin'
(ChronC_[Rositzke]:1011.18.1520)

c. **and** [$_{CP}$ Ælfmær abbud hi leton aweg. ]
and Ælfmær abbod they let away
'And abbot Ælfmær they allowed to get away.'
(ChronC_[Rositzke]:1011.18.1521)

d. **And** [$_{CP}$ hi ðær genamon inne ealle þa gehadodan men and weras and wif ]
and they there took inside all the ordained men and men and women
'And inside [the town] they took all the religious men and women.'
(ChronC_[Rositzke]:1011.21.1522)

It might be feasible to device more careful tests and inter-annotator agreement experiments to demonstrate systematic meaning differences between potential C-head conjunctions and logical connectors. However, this enterprise would go beyond the scope of this chapter and I will therefore merely mention it here as a possible future extension of this research.

**Wedged Conjunctions**

My final argument for the existence of C-head conjunctions is the following. The YCOE annotates several sentence tokens in such a way that the conjunction *and* is situated between an initial adverbial clause X and the rest of the main clause, `[X] and [IP ... ]`. These annotations could be incorrect or they could be indicative of disfluencies. However, they could also reflect the most plausible analysis of a sentence. In the latter case, the meaning of *and* must be paraphrased like that of one of its hypothesized adverbial ancestors and the conjunction can be parsed straightforwardly under `C`. A few examples are presented in (98).

The first sentence includes an initial *mid þy* 'when' - temporal adverbial clause. It is followed by *and* and the rest of the clause. The second part of the sentence cannot simply be analyzed as a second conjunct of the temporal clause, 'when ... and [when] ... ' because the sentence ends a section. The following sentence therefore starts a completely new idea. Hence, there remains no alternative main clause for the proposed coordinated subordinate clause to adjoin to. Instead, the following conjunction *and* could quite naturally mean 'then' and occur under `C`.

The second example involves a clause-initial *as*-clause, 'as my father knows me... .' The next clause cannot easily be parsed as a conjunct of the subordinate clause, 'as my father knows me and [as] I know him ...' because the subsequent material does not allow for a meaningful connection with such a constituent. Furthermore, the subsequent main clause also starts with a conjunction so that the the problem of creating an unusual position for a conjunction would simply be postponed to the next token. To avoid this problem, the conjunction *and* could be taken to mean 'so' or be semantically vacuous and be placed under `C`.

The third illustration shows a conditional *if*-clause before the conjunction *and* followed by another clause. Again, the second clause cannot easily be coordinated with the subordinate clause because of the context and because all the subsequent clauses also have initial *and*s until there is a section ending. Therefore, *and* in this case seems to be introducing the consequent of the conditional. It can conveniently be placed under `C`. Furthermore, a second conjunction, *oððe* 'or,' appears in clause-initial position. This structure may thus illustrate the co-occurrence of a logical connector and a C-head conjunction.

(98) a. [$_{CP}$ Mid þy we ða geornlicor þa stowe sceawodon & betwih þa bearwas eodon],
with that we then more-eagerly the place beheld and between the forests went
[$_{CP}$ [$_{C'}$ **&** [$_{IP}$ ic ða wynsumnesse & fægernesse þæs londes wundrade. ]]]
and I the loveliness and fairness of-the land wondered

'When we looked at the place more carefully and went between the forests, then I was amazed at the loveliness and fairness of the land'
(coalex,Alex:34.6.433)

b. [$_{CP}$ swa swa min fæder oncnæwð me.] [$_{CP}$ [$_{C'}$ **&** [$_{IP}$ ic oncnawe hine. ]]]
so as my father knows me and I know him

'As my father knows me, so I know him' (cocathom1,ÆCHom_I,_17:315.77.3154)

c. oððe [$_{CP}$ gif hit meteð ungesenodes mannes muð and lichoman]
or if it meets unblessed man's mouth and body
[$_{CP}$ [$_{C'}$ **and** [$_{IP}$ hit ðonne on forgietenan mannes innelfe gewiteð, ]]]
and it then in forgetful man's inside goes

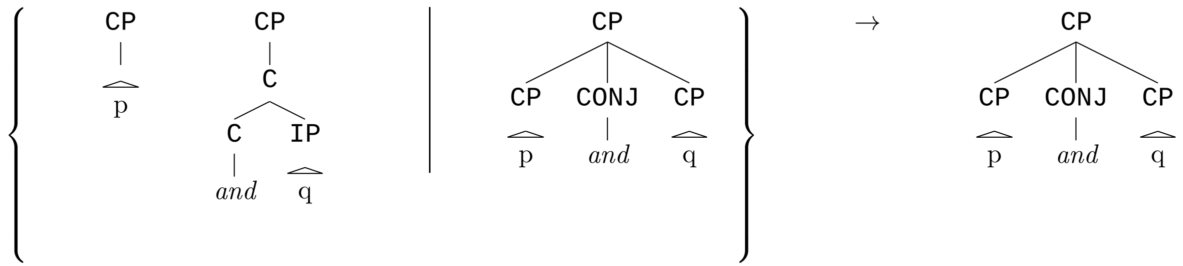
'Or, if he [=the devil] finds the mouth and body of an unblessed [unbaptized?] man, then he goes into this forgetful man's inside'
(cosolsat2,Sol_II:40.45)

Having proposed several arguments, albeit weak ones, for the existence of C-head conjunctions in Old English, I will now turn to their diachronic development. I will simply propose competition between C-head conjunctions as the conservative variant and logical connectors as the innovative variant. Eventually, C-head conjunctions disappear from the language. The graphic below represents the proposed change in terms of the general model of the implementation of a linguistic innovation in a group of speakers from the introduction chapter. It is not clear if there was ever a stage in the English language when conjunctions such as *and* categorically had to appear under `C` and thus I leave out the conservative stage of the change.

**Change in Coordination**

| | | |
|---|---|---|
| Conservative period: Speakers represent only the conservative variant $\alpha$ | Transitional period: Speakers represent both the conservative variant $\alpha$ and the innovative variant $\beta$ | Innovative period: Speakers represent only the innovative variant $\beta$ |
| $\alpha \rightarrow$ | $\{ \ \alpha \mid \beta \ \} \rightarrow$ | $\beta$ |
| Conservative period: Unknown if speakers ever represented conjunctions only as C-head conjunctions | Transitional period: Speakers represent conjunctions both as C-head conjunctions and as logical connectors | Innovative period: Speakers represent conjunctions only as logical connectors |

?



#### 3.3.1.9 Summary

In this section, I developed a computationally rigid grammar model of Old English clause structure.[3] It assigns to a large number of core structures of Early English a specific analysis. Admittedly, there remain a great many Old English syntactic patterns that my model will not be able to handle. However, the structures that I did manage to cover here are generated with an explicit and exhaustive algorithm and not just in an opaque or cursory way. This allows my model both to be used practically for research on Old English clause structure and to be falsified in a straightforward way.

One noteworthy feature of my model - as of any other grammar of Old English that I know of - is the generation of a great deal of structural ambiguity. For example, a simple 'subject - verb - object' sentence, such as *Petrus beheold Simon* 'Petrus beheld Simon' will receive no less than five parsing solutions under my model assumptions. In addition, this particular sentence allows for an additional analysis in which the first name is the object and the second the subject, resulting in the meaning 'Simon beheld Petrus.' The parsing solutions are enumerated in example (99) below.

(99) a. Petrus beheold Simon
Petrus beheld Simon
'Petrus beheld Simon' (coblick,LS_32_[PeterandPaul[BlHom_15]]:187.289.2415)

b. *Parse 1: Subject in Spec,IP, object in VP, I-initial parse*
[$_{\text{CP}}$ [$_{\text{IP}}$ Petrus [$_{\text{I}'}$ beheold [$_{\text{VP}}$ [$_{\text{V}'}$ Simon ]]]]]

c. *Parse 2: Subject in Spec,CP, object in VP, I-initial parse*
[$_{\text{CP}}$ Petrus [$_{\text{IP}}$ [$_{\text{I}'}$ beheold [$_{\text{VP}}$ [$_{\text{V}'}$ Simon ]]]]]

d. *Parse 3: Subject in Spec,IP, object extraposed, clausal spine terminates in I*
[$_{\text{CP}}$ [$_{\text{IP}}$ [$_{\text{IP}}$ Petrus [$_{\text{I}'}$ beheold ]] [$_{\text{DP}}$ Simon ]]]

e. *Parse 4: Subject in Spec,CP, object extraposed, clausal spine terminates in I*
[$_{\text{CP}}$ Petrus [$_{\text{IP}}$ [$_{\text{IP}}$ [$_{\text{I}'}$ beheold ]] [$_{\text{DP}}$ Simon ]]]

f. *Parse 5: Subject in Spec,VP, object extraposed, I-final parse*
[$_{\text{CP}}$ [$_{\text{IP}}$ [$_{\text{IP}}$ [$_{\text{I}'}$ [$_{\text{VP}}$ Petrus ] beheold ]] [$_{\text{DP}}$ Simon ]]]

g. *Parse 6: Object in Spec,CP, Subject in Spec,VP, I-initial parse*
[$_{\text{CP}}$ Petrus [$_{\text{IP}}$ [$_{\text{I}'}$ beheold [$_{\text{VP}}$ Simon ]]]]

However, I would claim that my model actually creates substantially less ambiguity than many of its competitors. Furthermore, my model at least makes it possible to list all the different analyses quickly and completely.

Two related points are worth mentioning briefly. First, the large amount of structural ambiguity is exploited and maybe sustained by information-structural considerations of Old English speakers. For instance, sentence elements in a high structural position tend to be given, functions lower in the structure tend to be new. Second, it seems plausible to assume that the massive structural ambiguity in Early English is reflective of many changes that are in progress at that time. As many competing grammatical options coincide and gradually alter the grammatical structure of the language, their optional application results in many alternative analyses for the same string.

My model has little to say about the first, information-structural point. The reason is that I do not currently have a good understanding of the way grammatical options, which are generated by the syntactic component, interface with speaker's usage choices, which are influenced by cognitive or functional processes. A crude treatment of information-structure with syntactic machinery, as features or structural positions, is not satisfactory for me. There are too many exceptions to rules associating a syntactic position with one particular discursive interpretation. I therefore merely included underspecified discourse functions in my model, which could perhaps serve as syntax-pragmatics interface points for future grammar engineering projects.

The second point, variable and changing competing rules, feature much more prominently in my model. To be precise, I included three syntactic changes: (1) Verb Positions become simplified such that verbs tend to occur more and more frequently under a head initial I-head. (2) Topicalization of an argument DP to the specifier of CP becomes less frequent overall. (3) C-head conjunctions are replaced by logical connectors. In the next section, I will describe how exactly the first two changes interact with the third.

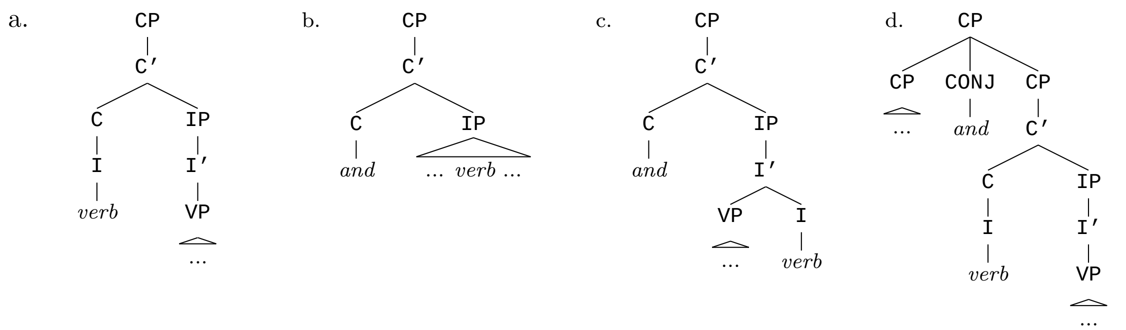
[3] The grammar model has been implemented as an LFG toy grammar that can be run on the Xerox Linguistic Environment (XLE) grammar engineering platform. This makes it easy to test the model predictions regarding the acceptance and rejection of certain strings and underlines its mathematical explicitness. The grammar can also be viewed with any word editor. It is located on the Dissertation DVD under `Chapter 3 - Conjunct Clause /01 Toy Grammar/Ch3ConjClause.lfg`.

### 3.3.2 Interactions between Old English Syntactic Phenomena

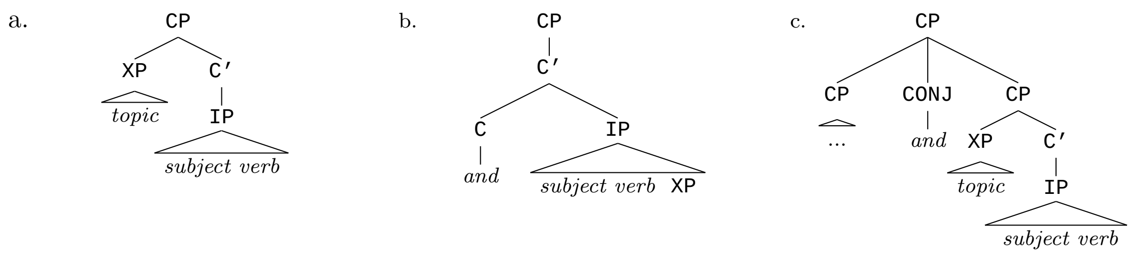
I will now explain how my model assumptions entail the differential syntactic behavior observed for MCs and CCs. The structural explanation, in turn, leads to a fair number of quantitatively testable predictions. In particular, my model predicts specific synchronic and diachronic interactions between changes in verb placement / topicalization and the emerging predominance of logical connector conjunctions. Finally, I will address the relevance of this research for historical linguistics in general.

#### 3.3.2.1 How the Model Explains the Special Status of Conjunct Clauses

CCs are more commonly verb-final than MCs. My analysis accounts for this fact in the following way. The verb can appear under `C` in contexts for high verb placement in all MCs, (100a). C-head conjunctions, however, may block the verb's appearance under `C`. They behave like complementizers in subordinate clauses in this respect. Consequently, C-head conjunctions will force higher rates of structures with the finite verb in `I` in CCs than MCs, (100b). As a side effect of this elevated overall rate of verbs within the `IP`, there will also be an increase in the number of verb-final patterns specifically, (100c). Underlyingly, the rate of `I`-final headedness is assumed to be identical in MCs and CCs. The distributional difference in verb placement between MCs and CCs follows largely from the same structural configurations as in den Besten's (1983) classical analysis of the verb-second constraint in Modern West Germanic. However, the complementarity between high verbs in `C` in contexts for high verb placement and the presence of conjunctions is merely a tendency and not categorical (unlike the complementarity between verbs in `C` and complementizers in subordinate clauses). The reason is that conjunctions are not necessarily inserted under `C` but may be logical connectors instead and can then co-occur with verbs in `C`, (100d).

(100)



CCs are also less likely to occur with a topicalized phrase. This can be explained as follows. My model allows MCs to freely apply the topicalization rule (at least as long as certain requirements imposed by information structure are obeyed), (101a). If a C-head conjunction is present, however, the application of the topicalization rule will be prohibited, (101b). However, this does not entail that the presence and absence of clause-initial topics distributes in a complementary way between MCs and CCs. Instead, the word order 'conjunction - topic - subject' can still be generated through the usage of logical connectors, (101c). Thus, the proposed analysis successfully models a reduction in, but not a complete lack of, Old English topicalized elements in CCs.

(101)



#### 3.3.2.2 Predictions about Interactions

The account of the distributional differences between MCs and CCs offered above leads to the following quantitatively testable predictions.

First, verbs should surface under `I` more commonly in CCs than MCs overall. The reason is the following. The total number of finite verbs in Old English can be regarded as the sum total of verbs under `C` and those under `I`. Hence, there is an equilibrium between high and low verb placement such that if the rate of the former

decreases, the rate of the latter increases and vice versa. Since CCs are assumed to lower the number of `V`-to-`C` clauses, they should surface with a higher rate of verbs under `I` in general and only as a side effect of this with a higher rate of verb-final patterns in particular. This hypothesis contrasts with the alternative assumption that it is only the number of `I`-final patterns specifically, and not the number of clauses with the verb within `IP` as such, that displays an elevated rate in CCs over MCs.
The hypothesis can be tested by counting and comparing the number of clauses with `V`-to-`C` patterns, `I`-final structures, and other cases of verbs in `IP`. Surface orders reflecting underlying `V`-to-`C` structure can be identified more easily with pronominal subjects than with full subjects. It therefore makes sense to collect the date separately for those two subject types. One should then observe a severe reduction in high verb placement under `C` in CCs by comparison with MCs. Furthermore, both verb-final and verb-medial structures should occur at elevated rates in CCs over MCs.

(102) **Hypothesis 1:** *Overall Larger Number of Verbs in `IP` in CCs than MCs*
A comparison between sentences with high verb placement, verb-final and verb-medial structures should reveal that, by comparison with MCs, CC display (i) a lower rate of high verbs and (ii) a higher rate of both verb-final and verb-medial structures.

Secondly, it should at least be possible for the underlying rates of `I`-final headedness and other alignments of `IP` to be identical in MCs and CCs. My model assumes that the actual rates of these phrase structures are the same in MCs and CCs. Surface distortions are only due to the availability of C-head conjunctions, which block verb placement under `C` and hence increase the number of verbs in `IP`. A statistical correction should therefore exist that distributes the instances of `V`-to-`C` sentences over verb-final and verb-medial clauses by the same measure in MCs and CCs and that consequently equalizes the rates of the variants of `IP`-headedness in the two clause types. If such a correcting factor does not exist, this would refute the assumption of identical rates of `IP` phrase structural configurations between MCs and CCs. If it does, this would make it at least plausible for the model assumption to be correct.
The question then becomes how one should determine the multiplier for the distribution of `V`-to-`C` sentences over clauses with verbs in `I`. One reasonable way would be to take the respective rates from subordinate clauses because their virtually categorical lack of verb placement under `C` yields relatively unbiased percentage estimates of structures with the verb in `I`. Hence, applying the relative frequency of `I`-final and potentially `I`-initial embedded clauses to the number of `V`-to-`C` structures in MCs and CCs should harmonize the behavior of those two clause types with respect to their surface realization of verb-final and verb-medial orders.

(103) **Hypothesis 2:** *Potential for Identical `IP`-Headedness in CCs and MCs*
If one distributes the instances of verbs placed under `C` in MCs and CCs over clauses with `I`-final phrase structure and other `IP`-alignments according to the rates of these two structural patterns found in subordinate clauses, then there should be a substantial convergence of the rates of `I`-final structures and other `IP`-alignments for these two clause types.

The third prediction concerns the diachronic development of high verb placement under `C`. Declaratives gradually lose high verb placement during the Old and early Middle English periods (see changes described under section 3.3.1.4). However, the change should proceed more slowly in CCs than in MCs if the assumption of the simultaneous loss of C-head conjunctions is correct. The reason is that the loss of conjunctions under `C` should free up this position for finite verbs, thus compensating for the general decline in high verb placement to some degree. Put differently, the disappearance of the Old English conjunct clause type runs orthogonal to the reduction in high verbs under `C` - hence the implementation of the former change should mitigate the surface effects of the latter. For the sake of concreteness, one could run the following simple simulation: Early Old English finite verbs might be placed under `C` 40% of the time overall, all things being equal, and 50% of all conjunctions of that period might be `C`-heads. Hence, MCs would realize 40 out of 100 finite verbs under `C`. CCs, on the other hand, would only show 20 out of 100 finite verbs under `C` since the rest, here 20 examples (50% of 40), would be blocked by C-head conjunctions. By late Old English, the probability of both high verb placement and C-head conjunctions might have halved. 20% of all finite verbs might now occur under `C` and 25% of all conjunctions might now be C-head conjunctions. Consequently, MCs would show 20 out of 100 finite verbs under `C`. High verb placement would have declined considerably, from 40 out of 100 to 20 out of 100 examples. CCs, in contrast, would still show 15 out of 100 verbs under `C` because only one quarter of all verbs, here 5 examples (25% of 20), would be blocked by C-head conjunctions. By comparison with MCs, high verb placement in CCs would thus have declined much less substantially, from 20 out of 100 to just 15 out of 100 examples. In other words, the loss of high verb placement would have proceeded more slowly in CCs than in MCs.

The predicted differential reduction in verb placement under C can be investigated by tracing the development of the data collected for the previous hypotheses across time. The data involving pronominal subjects are most appropriate because they reliably diagnose the specifier of IP. Verbs before such subjects indicate high verb placement under C, verbs after them a position inside the IP. A logistic regression that predicts verb positions from independent time and clause type variables should then show a significant effect of time on the loss of high verbs in general, and a significant interaction effect between time and clause type such that the decline happens more slowly in CCs in particular. Full subjects cannot easily be used for this measurement because they do not reliably diagnose the specifier of IP in Old English yet.

(104) **Hypothesis 3:** *Differential Development of V-to-C Contexts*
If one measures the proportion of 'verb - pronoun subjects' vs. 'pronoun subject - verb' structures over time, there should be a noticeable decline in the former word order option. This development should proceed more slowly in CCs than MCs.

Next, separating conjunctions and subjects from each other should have a substantial effect on the observed verb placement patterns. My formal model assumptions make it possible to classify observations of conjunctions into two groups, those that necessarily function as logical connectors and those that are at least potentially C-head conjunctions. Conjunctions must necessarily be logical connectors if they are removed from IP by another sentence element, [ conjunction ... X ... [$_{IP}$ ... ]]. They may be analyzed as C-heads if they are immediately adjacent to IP, [ conjunction [$_{IP}$ ... ]]. Necessarily logical connectors should not block placement of the verb under C and should thus not occur with an elevated rate of verb-final structures. Potential C-head conjunctions, on the other hand, should frequently prevent the verb from being placed under C and so it should be specifically those items that are responsible for the increased proportion of verb-final orders in CCs.
The hypothesis can be tested by dividing the CC data collected for the investigation of the first hypothesis into necessarily logical connectors and potential C-head conjunctions and subsequently observing their verb placement distributions. Since necessarily I-final structures can be recognized independently of subject type, the entire set of CCs, featuring both pronominal and full subjects, can be used for this study.

(105) **Hypothesis 4:** *Effect of Separation of Conjunction and IP*
If one divides all CCs into those clauses whose conjunction can potentially be parsed as a C-head and those whose conjunction must be a logical connector, the rate of verb-final structures should become noticeable greater in the former than the latter subtype.

The fifth hypothesis relates to the loss of I-final headedness (see changes described under section 3.3.1.3). The loss of verb-final structures should follow a rate of change that is different in clauses that may involve a C-head conjunction than in all other clause types. This can be explained as follows:
One must assume that the underlying rate of I-final headedness declines at the same speed in all contexts, in matrix as in embedded clauses. This is a reasonable assumption in light of the plausibility of the Constant Rate Hypothesis, as argued in chapter 2. Now, the loss of C-head conjunctions will allow more verbs to occur under C in CCs relative to other clause types over time. The additional verbs that are placed under C logically cannot occur inside IP at the same time and in an I-final configuration in particular. The decrease in conjunctions under C thus functions as an added cause for the loss of I-final patterns that is not shared by main clauses without conjunctions, or with logical connectors or with subordinate clauses. The loss of C-head conjunctions should therefore speed up the reduction in observed verb-final clauses over time. Metaphorically speaking, one might image C-head conjunctions as drain-stoppers. They are gradually pulled so that finite verbs get siphoned off from I-final position. A concrete simulation might help to further clarify this prediction. One could assume that early Old English has a 50% chance of realizing a conjunction as a C-head, a 40% probability for a finite verb to be placed under C, and 30% I-final headedness. 18 out of 100 finite verbs in MCs and 24 out of 100 CCs would then be expected to occur in an I-final configuration. By late Old English all of these probabilities should have dropped. For example, 30% of all verbs might now be placed under C, 20% of conjunctions might be C-heads, and 10% of all IPs might be head-final. In this case, there should now be 8 verb-final out of 100 total MCs, and 9 verb-final out of a total of 100 CCs. I-final structures would have declined by 10 percentage points in MCs but by a comparatively greater amount of 15 percentage points in CCs. In other words, CCs would have lost I-final structures more rapidly.
The prediction can be tested by regressing the data gathered for the preceding, fourth hypothesis diachronically against time. One should observe a significant effect of time on the loss of verb-final structures in general. Furthermore, the time - clause type interaction terms should not be significant for any pair of clause types except for the one involving CCs with potential C-head conjunctions. Here, the interaction term should be significant and suggest that these CCs change at a faster rate.

(106) **Hypothesis 5:** *Differential Development of* `I`*-Final Structures*
When the corpus texts' proportions of `I`-final structures are regressed against clause type and time, all clause types should show a constant rate effect for the decline in such structures except for clauses that may involve a C-head conjunction. These clauses should lose verb-final structures faster than the others.

Next, `I`-initial headedness increases during the Early English period (see the change described under section 3.3.1.3). The sixth prediction states that the rise of necessarily `I`-initial headedness should proceed at the same speed in both MCs and CCs. By necessarily `I`-initial clauses, I mean clauses that include a post-verbal non-postposing diagnostic element. In fact, there should not be a substantial difference even in the base probabilities to realize an `I`-initial clause between MCs and CCs. The reason for this is simply that the application of rules generating `I`-initial phrase structure happens independently of the presence or absence of C-head conjunctions. I will again show how exactly this prediction works with a simulation: Early Old English might have a 40% chance of placing a verb under `C`, 20% necessarily `I`-initial headedness out of all clauses with the verb in `I`, and a 50% probability for conjunctions to realize a `C`-head. In this case, 12 out of 100 MCs and 16 out of 100 CCs should surface with necessarily `I`-initial phrase structure, 48 out of 100 MCs and 64 out of 100 with other, potentially `I`-final headedness. If one now calculates the proportion of necessarily `I`-initial clauses out of all clauses with the verb under `I`, one finds a probability of 12 `I`-initial vs. 48 other `IP` clauses = 20% in MCs and 16 `I`-initial vs. 64 other `IP` clauses = 20% in CCs. The rate is equal to the stipulated input probability of necessarily `I`-initial clauses and completely independent of verb placement under `C` or its blockage by `C`-head conjunctions. By late Old English, verbs might only have a 20% chance of being placed under `C`, the rate of necessarily `I`-initial clauses out of all clauses with the verb in `I` might have increased to 40%, and the probability for a conjunction to be a C-head conjunction might have dropped to 25%. Under these assumptions, necessarily `I`-initial phrase structure would manifest in 32 out of 100 MCs and in 38 out of 100 CCs, other `IP` configurations in 48 out 100 MCs and 57 out of 100 CCs. A calculation of the percentage of necessarily `I` initial clauses out of all clauses with the verb in `IP` would show 32 `I`-initial vs. 48 other `IP` clauses = 40% in MCs and 38 `I`-initial vs. 57 other `IP` clauses = 40% in CCs. Again, the rate is identical to the stipulated input probability of necessarily `I`-initial phrase structure in both clause types; `V`-to-`C` structures or C-head conjunctions are irrelevant for the observed proportions. The percentage of necessarily `I`-initial clauses would have increased from 20% to 40% in MCs and likewise from 20% to 40% in CCs. Hence, the overall probability of finding a necessarily `I`-initial clause and the rate of change should be identical in the two clause types.
One can test this prediction relatively easily by carrying out a now commonly used measurement for the development of necessarily `I`-initial headedness, namely by comparing elements that cannot extrapose in post-verbal position to their occurrence in pre-verbal position. This yields a lower bound of necessarily `I`-initial clauses. These non-postposing diagnostics are (i) pronouns (possibly reinforced by *self*), (ii) particles, (iii) stranded prepositions and (iv) negative objects. The verb should follow an overt subject to make sure that the verb is placed under `I`. A logistic regression model that predicts the placement of diagnostics in post-verbal position from time and clause type should not significantly improve by adding the interaction term between these two variables.

(107) **Hypothesis 6:** *Uniform Development of Necessarily* `I`*-Initial Structures*
If the rise in `I`-initial headedness is measured with post-verbal diagnostics such as pronouns, particles, stranded prepositions and negatively quantified objects, one should observe a constant rate effect in the increase of such structures between MCs and CCs.

Hypothesis number seven concerns the development of object topicalization. Object topicalization decreases during the course of the Early English period (see the change described under section 3.3.1.6). The decrease should occur more rapidly in MCs than in CCs. The reason is that, as C-head conjunctions are lost, they no longer block the `CP`-domain for topicalization. The change in coordination thus runs orthogonal to the change in topicalization. The loss of C-head conjunctions should compensate for the decline in object topics to some degree. This in turn should slow down the decrease in object topicalization in CCs.
The hypothesis can be assessed as follows. Since pronominal subjects unambiguously identify the specifier of `IP`, objects that occur in front of them are very likely to have topicalized to the specifier of `CP`. Objects should be operationalized exclusively as full, nominal `DP`s. Fronted pronouns are ambiguous between topics and preposed elements in the `I`-domain. Other categories, notably `PP`s or `ADVP`s, probably occur as clausal adjuncts to `CP` rather than as topics in the specifier of `CP`. Such elements are therefore unsuitable for this study. One should then observe a larger number of objects in front of pronominal subjects in MCs than in CCs. A logistic regression model that predicts the occurrence of objects before subject pronouns from time and clause type should show an effect of time on the variation such that object fronting decreases, and a significant time - clause type interaction effect pointing towards a steeper decline in MCs than in CCs.

(108) **Hypothesis 7:** *Differential Development of Object Topicalization*
Measuring object topicalization as the occurrence of a nominal object `DP` before a subject pronoun, 'object - pronoun subject,' vs. its occurrence after a pronominal subject, 'pronoun subject ... object,' one should find (i) a greater frequency of fronted objects in MCs than in CCs, (ii) a substantial decline in the former word order option over time and (iii) a faster decline of this pattern in MCs than in CCs.

Finally, my eighth hypothesis predicts a differential development between MCs and CCs also for the topicalization rate of nominal subjects. The reasoning for this effect is identical to the logic behind the differential decline in object topicalization just described. However, it is impossible to measure directly the rate of topicalized subjects because 'subject - verb' sentences are ambiguous between parses with the subject in the canonical subject position, the specifier of `IP`, and the subject in the topic position, the specifier of `CP`. It is therefore necessary to measure the rate of subject topicalization indirectly.
One way of doing this reads as follows: One can identify structures in which subject topicalization is a potential option and compare them to structures in which subject topicalization cannot possibly have taken place. Preposed non-subject pronouns can be used for this purpose. The pattern 'full subject - pronoun - verb' is not unlikely to involve a preposed pronoun inside the `I`-domain and a topicalized subject, although other parses remain available. The pattern 'pronoun - full subject - verb,' in contrast, is entirely incompatible with parses that involve the subject topicalization rule. One should then observe fewer structures with potential subject topicalization in CCs than in MCs. Furthermore, as time goes by, the rate of potential subject topicalization should harmonize between the two clause types. How exactly this convergences should occur is hard to predict. However, a logistic regression model should, once again, show a significant time - clause type interaction term, which points towards a significantly divergent development in subject topicalization between MCs and CCs. By the end of the Early English period, the base probabilities of finding a potentially topicalized subject in MCs and CCs should have become substantially more similar, maybe even roughly identical.

(109) **Hypothesis 8:** *Differential Development of Subject Topicalization*
If one measures clauses in which subject topicalization may have taken place as 'subject - pronoun' patterns in comparison to clauses in which subject topicalization cannot possible have taken place as 'pronoun - subject' orders, one should (i) observe a reduced potential for subject topicalization in CCs by comparison with MCs and (ii) a differential development of potential subject topicalization in MCs and CCs such that the clause type difference diminishes over time.

#### 3.3.2.3 Relevance of the Interactions

The investigations of the eight distributional hypotheses listed in the last section can, I believe, greatly contribute to the development of good methodologies and to the generation of novel insights in the field of historical linguistics.

Firstly, hypothesis 3 regarding `V`-to-`C` patterns, hypothesis 5 about `I`-final headedness, and hypotheses 7 and 8 concerning object and subject topicalization respectively are the first, or among the first, <u>theoretically predicted violations of constant rate effects</u>. These violations do not arise because the Constant Rate Hypothesis is supposedly wrong but because several syntactic changes interact in such a way that one apparently slows down or speeds up the other. In fact, the hypotheses rely on the correctness of the Constant Rate Hypothesis as a default. If these hypotheses can be supported by evidence, this would therefore constitute a new way of demonstrating the correctness of the Constant Rate Hypothesis as well.

Secondly, the proposed investigation of interactions between syntactic changes may be a relatively <u>new approach to the study of diachronic variation</u>. Specifically, if the hypotheses turn out to be acceptable, they would show that syntactic changes can influence each other, and that the direction of the interactions could be predicted and measured in corpora. One of the changes involved in the present study, the loss of C-head conjunctions, is measurable only in an indirect way and occurred during a period whose surviving text material is relatively limited and deficient. Future case studies could perhaps identify interactions between syntactic changes that are more straightforward to operationalize and that could be measured in a richer corpus. The possibility of multiple syntactic changes that run orthogonal to each other should also be considered as a possible confounding factor when measuring changes in the future.

Apart from these more abstract relevancies, the investigation of the distributional hypotheses may also lead to a <u>better understanding of the special status of Old English conjunct clauses</u>. If the interactions between Old English conjunctions and verb placement or conjunctions and topicalization do indeed manifest as predicted, this would strongly suggest that Old English conjunctions do indeed disable parts of the left periphery as a syntactically active part of the clause and that this is the origin of the peculiar syntactic behavior of Old English coordinated main clauses.

## 3.4 Measuring Interactions between Old English Syntactic Phenomena

In this second, empirical part of the chapter, I will test eight predictions about Early English interactions between the presence of conjunctions and verb placement and the presence of conjunctions and topicalization. These hypotheses were derived from my theoretical Old English grammar modeled developed in the last part. Thus, support for their correctness should also strengthen one's confidence in this model. Furthermore, since a fair number of the predictions are diachronic in nature, their investigation will also inform this chapter's general research question of whether theoretically predicted interactions between syntactic changes can result in empirically measurable divergences in corpus data. I will first outline how I prepared the data and define certain syntactic notions. After this, I will sequentially test each of the eight distributional hypotheses. Finally, I will summarize my findings.

### 3.4.1 Data Preparation

Before I executed search queries on the electronic text files, I conducted several revision queries in order to recode certain aspects of the annotation scheme. First, I prefixed the label `SUBJECT-` to all referential or expletive, resumptive or non-resumptive, extracted or local non-predicative nominative noun phrases. This comprises the set of labels `NP-NOM|NP-NOM-x|NP-NOM-RSP|NP-NOM-#| NP-NOM-x-#|NP-NOM-RSP-#` for the Old English and `NP-SBJ|NP-SBJ-#|NP-SBJ-RSP` for the Middle English corpus files. Secondly, I added the label `PRONOUN-` to the beginning of all remaining `NP`s that immediately dominated a pronoun, `PRO|PRO^*`. This means that rare instances of complex pronouns, such as *them all*, and pronouns reinforced by self, such as *myself*, were included in this group as well. Finally, I marked adverb phrases, `ADVP*`, as `LIGHTADVERB-` if they only contained a single adverb.[4]

### 3.4.2 Definition of Syntactic Concepts

I defined a number of syntactic concepts for the subsequent hypothesis tests. First, I followed the corpus annotations for the definition of declarative main clauses, `IP-MAT*`, or subordinate clauses, `IP-SUB*` contained in a `CP` with a `C` node. I defined the following main clause subtypes. Ordinary main clauses, MCs, were any kind of main clause that did not contain a conjunction (or Old English negative conjunction), `NEG+CONJ|CONJ`. Coordinated main clauses, CCs, in contrast, were declarative matrix clauses that did contain such an element.

Next, I distinguished between two kinds of subjects. Pronominal subjects were subjects immediately dominating only a nominative pronoun, `PRO^N`, the indefinite pronoun *man* 'one', `MAN*`, or a nominative demonstrative, `D^N`. The inclusion of the last part-of-speech label may not conform well to the intuitive sense of 'pronominal subject' for some researchers. Demonstratives were included here because, like the other categories, they have a very low probability of occurring in the lower subject position. Full or nominal subjects were subject labels that immediately dominated a common or proper noun head, a quantifier or a numeral, `N^N|NR^N|*Q|*Q^N|NUM*` for Old English texts, and the closely related labels `N|NS|*NPR*|*+N*Q|*ONE|NUM*` for Middle English texts.

Third, finite verbs comprised all labels for a finite verb, both lexical and pre-auxiliary, for Old English, labels with and without contracted negation or particles, and all moods, imperatives, indicative and subjunctive. Similarly, non-finite verbs were defined as the set of all infinitival and participle labels or non-finite clauses, `IP-INF*`. The list of the relevant verbal part-of-speech labels is quite long and so I will omit it here. The labels can be looked up easily in the online corpus documentations.

Fourth, I defined a class of heavy elements, abbreviated 'X,' as sentence elements that were required to be present in a clause for some investigations. This X group comprised the set of the following functions: (i) any noun phrase, `NP*`, which, after the sorting of pronouns explained earlier, were non-pronominal, (ii) other functions, namely adjectival phrases, `ADJP*`, small clauses, `IP-SMC*`, adverb phrases, `ADVP*`, whose one-word members had been relabeled earlier and so were relatively heavy, particles, quantifier phrases and stranded prepositions, (iii) prepositional phrases, `PP`, co-occuring with any additional constituent, which reduced the probability of the `PP` functioning as an adjunct, or (iv) any three phrasal categories.

Finally, by 'nominal object' I mean `NP`-nodes with or without case extensions and with or without a hashtag indicating displacement for Old English, first and second objects and `NP`s with a displacement hashtag for Middle English. Again, these specific `NP`s will not be pronominal because pronouns were coded separately earlier. I did not restrict the head or the length of the object in any way but merely required that it should be overt.

[4] The revision queries can be found on the Dissertation DVD under `Chapter 3 - Conjunct Clauses/02 Corpus/Revision Queries`.

### 3.4.3 Hypothesis Testing

I will now test the hypotheses deduced from my grammar model in section 3.3.2.2. In total, there are eight predictions concerning synchronic distributional differences between MCs and CCs and divergent diachronic patterns between these two clause types.

#### 3.4.3.1 Hypothesis 1: Overall Larger Number of Verbs in `IP` in CCs than MCs

My model predicts that C-head conjunctions should reduce constructions with high verb placement under `C` and therefore surface with a higher rate of clauses with the finite verb inside the `IP`. Hence, both `I`-initial and `I`-final structures should be more common in CCs than MCs. This prediction can be tested in a straightforward way by comparing the number of sentences with the verb under `C`, the verb in `I`-final structures and the verb under `I` in other configurations for the two different clause types.

**Data Collection**

I counted the incidence of three word order types in MCs and CCs. The structures were collected separately for pronominal and full subjects as well as for clauses with single finite verbs and clauses with additional non-finite verbs.

I will first describe the three word order patterns with pronominal subjects. Subject pronouns are unvaryingly placed in the specifier of `IP`. Thus, they reliably indicate the `CP`-`IP`-boundary.

(1) **High verb placement / `V-to-C`**. The first word order type shows the finite verb immediately preceding the pronominal subject. This *VS...* pattern necessarily diagnoses instances of clauses that are parsed with the rule for high verb placement under `C`, (39). Examples are shown for finite lexical verbs in (110) and for non-finite verbs in (111). Here and subsequently, the a.-examples illustrate for MCs, the b.-examples for CCs.

(110) **High verb placement / `V-to-C`:**

a. *V - Pronominal subject*

[$_{\texttt{CP}}$ þonne *geleanað* [$_{\texttt{IP}}$ he hit us swa us leofast bið ]]
then repays he it us as us dearest is

'Then he will repay us in a way that will be most pleasing to us' (cowulf,WHom_2:71.47)

b. *Conjunction ... V - Pronominal subject*

& [$_{\texttt{CP}}$ þa *æteawde* [$_{\texttt{IP}}$ he hine Nerone þy þriddan dæge ]]
and then appeared he reflexive to-Nero the third day

'And then he appeared to Nero on the third day'
(coblick,LS_32_[PeterandPaul[BlHom_15]]:183.221.2364)

(111) **High verb placement / `V-to-C`:**

a. *V - Pronominal subject*

[$_{\texttt{CP}}$ *ne beo* [$_{\texttt{IP}}$ ðu ðeah to upahafen ]]
not be you though too up-heaved

'Don't be too proud, though' (cocura,CP:17.119.21.802)

b. *Conjunction ... V - Pronominal subject*

& [$_{\texttt{CP}}$ *ne beo* [$_{\texttt{IP}}$ ge upahafene ]]
and not be you up-heaved

'And don't be proud' (cowsgosp,Lk_[WSCp]:12.29.4691)

The other two word orders show the finite verb after the pronominal subject. Such 'subject ... verb' sentences must contain the verb somewhere within the `IP`.

(2) **Verb-medial / Potentially `I`-initial**. The second word order type involves a subject pronoun, immediately followed by the finite verb, and a subsequent member of the group of heavy constituents, X. Such *SVX* sentences can potentially be parsed with the rule for head-initial `IP`s, (17a). However, verb-medial structures of this kind typically permit an alternative parse with postposition of the heavy constituent, e.g., (63), and the rule for termination of the clausal hierarchy in `I`, (17b). The examples in (112) illustrate.

(112) **Verb-medial / Potentially I-initial (I-initial or postposition):**

a. *Pronominal subject - V ... X*

[$_{\text{IP}}$ hi *underfengon* edlean heora weorca ]
they received reward of-their works

'They received the reward for their works' (cocathom1,ÆCHom_I,_28:416.170.5564)

b. *Conjunction ... Pronominal subject - V ... X*

**&** [$_{\text{IP}}$ hi *underfengon* þone halgan gast ]
and they received the holy ghost

'And they received the Holy Ghost' (cocathom1,ÆCHom_I,_22:357.86.4389)

Only a relatively small number of *SVX* sentences must necessarily be parsed with the rule for I-initial headedness, (17a). This is the case if a non-postposing diagnostic element is positioned after the verb, such as a pronoun or a particle etc. Examples with post-verbal pronouns are shown in (113). These sentences were picked up by my search queries because they involve at least three phrases of any kind after the finite verb.

(113) **Verb-medial / Potentially I-initial (necessarily I-initial):**

a. *Pronominal subject - V ... Pronoun ... X*

[$_{\text{IP}}$ He *gesette* **hine** to ealdre ofer an werod ]
he set him as leader over a troop

'He made him the leader of a troop' (coaelive,ÆLS_[Sebastian]:13.1218)

b. *Conjunction ... Pronominal subject - V ... Pronoun ... X*

**And** [$_{\text{IP}}$ he *hwyrfþ* **hine** þonne on þa wynstran ]
and he turns himself then on the left

'And then he will turn to the ones on the left' (conicodD,Nic_[D]:91.83)

I allowed pronouns or light adverbs to intervene between the subject and the finite verb for this context. If a verb-medial clause features such a light element, its structural ambiguity increases further. It could still be analyzed as an I-initial clause with high adjunction of the adverb (not implemented) or a rule for preposing of the pronoun, (79), and the rule for I-initial headedness, (17a), or with the same analysis for the adverb or pronoun, postposition of the post-verbal sentence element(s), e.g., by rule (63), and the rule for termination of the clausal hierarchy at I, (17b), or with postposition of the post-verbal sentence element(s), the light element inside a pre-verbal VP and the rule creating I-final headedness (25). Relevant examples are given in (114).

(114) **Verb-medial / Potentially I-initial (I-initial or postposition or I-final):**

a. *Pronominal subject - Pronoun - V ... X*

[$_{\text{IP}}$ He **him** *sealde* ricu oðerra kynrena ]
he them gave empires of-other nations

'He gave them empires of other nations' (cocura,CP:50.391.2.2650)

b. *Conjunction ... Pronominal subject - Light adverb - V ... X*

**And** [$_{\text{IP}}$ hi **þa** *wuldrodon* þa anwaldan and hergendlican þrynysse ]
and they then praised the powerful and excellent three-ness

'And they then praised the powerful and excellent Trinity' (coeust,LS_8_[Eust]:451.471)

For searches with non-finite verbs, I did not require the presence of a heavy constituent X, but replaced it with a non-finite verb. Hence, I looked for the pattern '(conjunction ...) subject - finite verb ... non-finite verb.' Again, pronouns or light adverbs were allowed to intervene between subject and finite verb. Examples are shown below.

(115) **Verb-medial / Potentially `I`-initial (`I`-initial or postposition of `I`-final):**

a. *Pronominal subject - $V_{pre-aux}$ ... $V_{non-fin}$*

[$_{IP}$ Ic *wille* sendan flod ofer ealne middaneard ]
I will send flood over all middle-earth

'I will send a flood over the whole world' (cocathom1,ÆCHom_I,_1:185.189.194)

b. *Conjunction ... Pronominal subject - Pronoun/ Light adverb - V ... X*

**Ac** [$_{IP}$ ic þe nu *wille* secgan hwilc se læacecræft is ]
but I you now will say which the leech-craft is

'But I will tell you now what that medicine is' (coboeth,Bo:22.51.1.926)

(3) **Verb-final / necessarily `I`-final**. Lastly, I collected all main and coordinated main clauses that have a heavy element X intervening between the pronominal subject and the finite verb. Instances of this *SXV* pattern can only be parsed by employing the rule for `I`-final phrase structure in (25). Verb-final structures of this kind are exemplified in (116)

(116) **Verb-final / necessarily `I`-final:**

a. *Pronominal subject ... X ... V*

[$_{IP}$ He þa forðon Drihtnes willan *sohte* ]
he then therefore Lord's will sought

'Therefore, he then sought the Lord's will' (coblick,LS_17.1_[MartinMor[BlHom_17]]:225.265.2900)

b. *Conjunction ... Pronominal subject ... X ... V*

**&** [$_{IP}$ he ær þone feþan *sohte* ]
and he earlier the warfare sought

'And earlier he sought battle' (coblick,LS_12_[NatJnBapt[BlHom_14]]:167.105.2115)

For searches with non-finite verbs, I replaced the heavy constituent X with a search for the non-finite verb itself. Hence, I retrieved structures of the type '(conjunction...) subject pronoun ... non-finite verb ... finite verb.' Examples are given in (117).

(117) **Verb-final / necessarily `I`-final:**

a. *Pronominal subject ... $V_{non-fin}$ ... $V_{pre-aux}$*

swa [$_{IP}$ he ma monna oleccan *sceal* ]
so he more men flatter shall

'In this way, he must flatter more people' (coboeth,Bo:26.60.9.1115)

b. *Conjunction ... Pronominal subject ... $V_{non-fin}$ ... $V_{pre-aux}$*

**ac** [$_{IP}$ ic oþere anginnan *sceal* ]
but I other begin shall

'But I shall begin another one [=another book]' (coorosiu,Or_2:8.53.4.1017)

Next, I counted the same three word order patterns with full subjects. Full subjects can be positioned more variably than pronominal subjects. Consequently, they do not indicate the intended structures as reliably. In particular, the first *VS* surface word order may be parsed in several ways. On the one hand, it is probably indicative of the desired structure with the verb under `C` when the verb appears in the same `V`-to-`C` contexts that pronoun subjects are found in, i.e., after an operator adverb, with negation or non-indicative mood, (118). On the other hand, the same pattern is often likely to reflect an undesired structure with the finite verb in `I` and the full subject in the low subject position. This is the case in verb-first, (119), or topic-initial main clauses, (120). There are thus many instances of *VS* patterns that do not diagnose high verb placement under `C` with full subjects so that their sum total is expected to be inflated relative to the counts with pronominal subjects. The first two examples illustrate for clauses with finite main verbs, the last for clauses that also contain a non-finite verb.

(118) **High verb placement (Probably V-to-C):**

a. *V - Full subject*

[$_{CP}$ þonne *gangað* [$_{IP}$ þa seofon steorran on uhtan upp ]]
then go the seven stars on evening up

'Then, the seven stars come up in the evening' (comart3,Mart_5_[Kotzor]:My9,A.3.786)

b. *Conjunction ... V - Full subject*

& [$_{CP}$ *ne forsio* [$_{IP}$ nan mon ðine gioguðe. ]]
and not scorn.subjunctive no man your youth

'{Let no man / May no man / No man should } scorn your youth' (cocura,CP:49.385.29.2612)

(119) **High verb placement (Possibly not V-to-C):**

a. *V - Full subject*

[$_{IP}$ *Wenað* [$_{VP}$ unwise læcas þæt þæt sie lendenadl oððe miltewærc ]]
think unwise physicians that that be loin-disease or spleen-pain

'Unwise physicians think that that's a disease of the loins or a pain in the spleen'
(colaece,Lch_II_[2]:31.1.13.2723)

b. *Conjunction ... V - Full subject*

& [$_{IP}$ *wenað* [$_{VP}$ menn ðæt he hit do for kystum ]]
and think menn that he it do for virtue

'And men think that he does it for virtue' (cocura,CP:20.149.6.1009)

(120) **High verb placement (Very probably not V-to-C):**

a. *V - Full subject*

[$_{CP}$ Under Moyses æ [$_{IP}$ *moste* [$_{VP}$ se bisceop habban an geæwnod wif ]]]
under Moses' law must the bishop have an espoused wife

'Under Moses' law, the bishop must have an espoused wife' (coaelive,ÆLS[Peter's_Chair]:218.2418)

b. *Conjunction ... V - Full subject*

**ond** [$_{CP}$ of þæam [$_{IP}$ *sceoldan* [$_{VP}$ regnas ofer eorðan cuman ]]]
and from that should rains over earth come

'And from that [=the sea], rain should come over the earth'
(comart3,Mart_5_[Kotzor]:Ma20,A.2.370)

The other two structures with the finite verb inside the IP behave alike with pronominal and with full subjects for my purposes. While my model adds parsing complexity in the form of the potential for subject topicalization to the specifier of CP, verb-medial *SVX* structures with full subjects can still potentially be parsed with I-initial headedness. A typical example is shown in (121). An additional example with a pronoun intervening between subject and verb is given in (122). The rudimentary labeled bracketing indicates the desired, potentially I-initial analysis.

(121) **Verb-medial / Potentially I-initial (I-initial or postposition):**

a. *Subject - V ... X*

þæt word [$_{I'}$ *is* ælmihtig God ]
that word is almighty God

'That word is the almighty God' (cocathom1,ÆCHom_I,_25:385.189.4940)

b. *Conjunction ... Subject - V ... X*

& þæt word [$_{I'}$ *wæs* God ]
and that word was God

'And the word was God' (cocathom1,ÆCHom_I,_2:195.166.421)

(122) **Verb-medial / Potentially I-initial (I-initial or postposition or I-final):**

a. *Subject - Pronoun - V*

Ærryst se casere him [$_{I'}$ *bead* gold ond seolfor ]
first the emperor them offered gold and silver

'First, the emperor offered them gold and silver' (comart3,Mart_5_[Kotzor]:Se24,B.8.1838)

b. *Conjunction ... Subject - Pronoun - V*

**ond** se casere him [$_{I'}$ *bead* gold ond seolfor ]
and the emperor him offered gold and silver

'And the emperor offered him gold and silver' (comart3,Mart_5_[Kotzor]:Ju15,A.11.963)

Similarly, the pattern *SXV*, including patterns of the kind 'full subject ... non-finite verb ... finite verb,' reliably indicates I-final headedness with full subjects. The sentences in (123) illustrate.

(123) **Verb-final / necessarily I-final:**

a. *Subject ... X ... V*

Moyses [$_{I'}$ ða Godes hæse *gefylde* ]
Moses then God's order fulfilled

'Then Moses fulfilled God's order' (cocathom2,ÆCHom_II,_12.1:114.133.2471)

b. *Conjunction ... Subject ... X ... V*

**and** his geferan [$_{I'}$ ða mid fleame *ætburston* ]
and his companions then with flight away-burst

'And his companions then escaped by flight' (cocathom2,ÆACHom_II,_14.1:141.114.3124)

I conducted a total of 3 (word order patterns) * 2 (clause type) * 2 (subject type) * 2 (finite vs. non-finite verb) * 2 (period; Old and Middle English) = 48 search queries.[5]

**Results**

I will first evaluate the results for clauses with subject pronouns. Pronominal subjects are reliable diagnostics of the specifier of IP. Hence, the three word orders *VS*, *SXV* and *SVX* (where X can also be a non-finite verb) directly correspond to verb placement under C, in necessarily I-final and potentially I-initial position respectively. The relevant numbers are presented in table 3.4.

| Clause Type | *V S* High verb V-to-C | *SVX* Verb-medial Potentially I-initial | *SXV* Verb-final Necessarily I-final |
|---|---|---|---|
| MC | 7,629 (41.2%) | 10,125 (54.7%) | 766 (4.1%) |
| CC | 1,528 (15.7%) | 6,484 (66.7%) | 1,716 (17.6%) |

Table 3.4: Verb positions in MCs and CCs with pronominal subjects

41.2% of all sampled MCs place the verb before a subject pronoun and hence under C, while only 15.7% of all CCs do so. This difference is significant ($\chi^2$=1889.8, df=1, $p$<0.001). Furthermore, the clause type effect is quite strong, as shown by the fact that the odds of finding a V-to-C-sentence (vs. a sentence with the verb under I) are more than 3 times greater in MCs than CCs (odds ratio = 3.76, 95%-confidence interval [3.53 – 4.00]). This confirms that CCs display a reduced rate of high verb placement. Next, MCs are I-final 4.1% of the time and show some other IP-configuration in 54.7% of all cases. In contrast, the proportion of I-final clauses is much higher in CCs at 17.6%, and, crucially, so is the percentage of structures with potentially I-initial alignments at 66.7%. This suggests that CCs do indeed manifest more structures with the verb within the IP overall and, possibly only as an indirect consequence of this, a greater incidence of verb-final structures specifically.

[5]The CorpusSearch query and output files can be found on the Dissertation DVD under `Chapter 3 - Conjunct Clauses/03 Hypothesis Testing/H1`.

I will now turn to the three word order patterns with full subjects. The relevant data is shown in table 3.5.

| Clause Type | *V S* (structurally ambiguous) | *SVX* Verb-medial Potentially `I`-initial | *SXV* Verb-final Necessarily `I`-final |
|---|---|---|---|
| MC | 9,706 (55.6%) | 7,010 (40.1%) | 751 (4.3%) |
| CC | 2,358 (35.3%) | 3,496 (52.3%) | 825 (12.4%) |

Table 3.5: Verb positions in MCs and CCs with full subjects

There appear to be more *VS* sentences with full than with pronominal subjects (55.6% vs. 41.2% for MCs, 35.3% vs. 15.7% for CCs). This surplus is expected as a consequence of the large number of *VS*-structures that do not reflect instances of the targeted `V`-to-`C` construction but parses with the finite verb in `I` and the subject in a lower subject position.

Nevertheless, the data shows the same trends as before: High verb placement is more common in MCs than CCs at 55.6% vs. 35.3% respectively. The difference is significant and the size of the effect is moderately large ($\chi^2$=792.7, df=1, $p$<0.001, odds ratio = 2.29, 95%-confidence interval [2.16 – 2.43]). CCs surpass MCs in the rate of verbs under `I` both in verb-final clauses (12.4% for CCs vs. 4.3% for MCs), and, importantly, for verb-medial clauses as well (52.3% for CCs vs. 40.1% for MCs). This finding points, again, to an increased rate of structures with the verb under `I` for CCs overall causing a greater incidence of verb-final structures in this clause type as a side effect.

Bech also recognizes the balance between the increased frequency of *SXV* and the lowered frequency of *(X)VS* conjunct clauses (2001: e.g., 192-5). However, she does not recognize the potentially causal relation between these two observations and merely seeks functional explanations for them individually. The explanation for the reduction in high verb placement, for instance, reads, "XVS word order is particularly suitable for introducing new turns of events and heavy, new, high I[nformation]V[alue] subjects. [...] [I]t is hardly surprising that the XVS pattern does not contain very many conjunct clauses, since [this function is not] compatible with the function of conjunct clauses to any great extent" (ibid.: 189). This does not seem to me to be a proper explanation because the explanandum ("Why do conjunct clauses appear less with *XVS* and their associated functions?") is simply restated in the explanans ("Conjunct clauses are those clauses that are incompatible with *XVS* and their associated functions."). Furthermore, the reference to an initial constituent X is both unhelpful as it subsumes fundamentally different elements, i.e., operator adverbs, sentence adverbs and argument topics, as well as irrelevant for the distribution of finite verbs. I certainly acknowledge the existence of (not categorical but probabilistic) correlations between word order patterns and functional factors. However, functional factors should be seen as secondary, as consequences or information-structural exploitations of the availability of syntactic resources, not as primary, as an antecedent cause or explanation for their existence. My model therefore tries to account for the reduced number of *VS* patterns in CCs as a part of a wider prediction of verb placement dependencies from a syntactic mechanisms that partially prevents verbs from being re-written under `C`, i.e., the presence of C-head conjunctions.

All in all, the data supports the first hypothesis. Verbs are less likely to appear under `C` in CCs than MCs. In contrast, the former clause type shows an elevated rate of verbs under `I` in general and consequently a higher rate of `I`-final configurations in particular. My model explains this as a consequence of the presence of C-head conjunctions. I acknowledge that the word order patterns correlate with particular information structural values of the sentence elements it contains, but do not formalize them as an explanatory variable.

#### 3.4.3.2 Hypothesis 2: Potential for Identical `IP`-Headedness in CCs and MCs

My model predicts that the rates of `I`-initial and `I`-final phrase structures should be identical in MCs and CCs. Differences in the observable realization of these patterns are supposed to be due solely to an elevated rate of high verb placement in MCs. Unfortunately, there does not seem to be a way to establish that these rates do not differ substantially between MCs and CCs in a technical and unequivocal manner. However, I should at least be able to demonstrate the possibility of identity in the underlying rates of verb-final and verb-medial structures in those two clause types. I will test this prediction by calculating the relative frequencies of verb-final and verb-medial structures in subordinate clauses and by using the obtained estimates to assign the observed high verb placement structures in MCs and CCs to the verb-final and verb-medial categories instead.

### Data Collection

I collected the same three word order patterns as for the previous hypothesis in subordinate clauses, SCs, to calculate an unbiased measure of the rates of verb-final and verb-medial phrase structure. If the subordinate clause was conjoined, only the first conjunct was considered.

First, I counted the number of structures with a finite verb in front of the subject. This *VS* pattern is expected to be rare in subordinate clauses because the complementizer should block verb placement under C. Where *VS* does occur, it most likely does not reflect an embedded clauses with verb placement under C but some other, irrelevant parse. For instance, a subject may invert with the verb after a subordinator because CP-recursion allows for main clause syntax or because a performance error leads to a disfluency. Such cases are exemplified for pronominal subjects in (124a) and for full subjects in (124b).

(124) **High verb placement (Probably CP-recursion or disfluency):**

a. *V - Pronominal subject*

ac ic secge him swyðor soðlice Godes word, [$_{CP}$ þæt
but I say him rather truly God's word that
[$_{CP}$ se þe lufað on eorþan his eorðlican fæder, modor, oððe bearn, oþþe wif ofer God
he who loves on earth his earthly father, mother or child, or wife over God,
[$_{C'}$ *ne bið* [$_{IP}$ he Gode wurð. ]]]]
not is he to-God worthy

'But rather I will truly tell him God's word [namely] that, as for him who loves his earthly father, mother, child or wife more than God [here] on earth, this person is not worthwhile for God' (coaelive,ÆLS_[Thomas]:375.7781)

b. ... [$_{CP}$ forðon þe [$_{CP}$ his sweostor, seo wæs Osþryð haten, [$_{IP}$ *hæfde* Æðelred Mercna cyning him to wife ]]]
... because his sister who was Osthryth called had Athelred, Mercians' kind himself to wife

'... because Athelred, king of Merica, took his sister, who was called Osthryth, as a wife for himself' (cobede,Bede_4:22.324.17.3254)

SCs with full subjects may also show post-verbal subjects because they may be placed in the lower subject position. Therefore, the number of embedded VS sentences is expected to be higher with full than with pronominal subjects. Illustrations are shown in (125).

(125) **High verb placement (Probably low subject):**

a. on ðam anum wæs corn swa hit gecweden is. swa fela [$_{CP}$ swa [$_{IP}$ *bið* [$_{VP}$ sandceosol on sæ ]]]
in that one was corn so it said is so much so is sand-pebble in sea

'In that one [country, Egypt], there was corn, like it is said, as much as there is sand in the sea.' (cocathom2,ÆCHom_II,_12.1:111.30.2391)

b. ... [$_{CP}$ þæt [$_{IP}$ him *gewearð* [$_{VP}$ se þridda penig of þære tolne on Sandwic ]]]
... that him became the third penny of the toll at Sandwich

'... that he obtained the third penny of the toll at Sandwich' (codocu3,Ch_1467_[Rob_91]:6.154)

c. ... on þam stede [$_{CP}$ þe [$_{IP}$ *lagon* [$_{VP}$ þa twegen gebroþra bebyrigde on ær ]]]
... in the place that lay the two brothers buried in earlier

' ... in the place that the two brothers lay buried in earlier' (coaelive,ÆLS_[Cecilia]:282.7278)

The pattern *SVX*, in contrast, is extremely numerous in SCs and likely to reliably reflect potentially I-initial structures, (126). I illustrate with a full subject below. As before, only pronouns or light adverbs were allowed to intervene between subject and verb.

(126) **Verb-medial / Potentially I-initial (I-initial or postposition):**
*Subject - V ... X*

... [$_{CP}$ oððæt [$_{IP}$ Aaron *eode* [ut of ðam getelde] [mid his storcyllan] ]]
... until Aaron went out of the tent with his incense-holder

'... until Aaron went out of the tent with his censer' (coaelhom,ÆHom_21:261.3213)

Likewise, the word order *SXV* is common in SCs and assumed to diagnose necessarily `IP`-final structures, (127). The example below illustrates a case with a full subject. It was retrieved because at least three constituents, or a `PP` and an additional phrase, intervene between subject and finite verb.

(127) **Verb-final / necessarily `I`-final:**
*Subject ... X ... V*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ... þa | [$_{\texttt{CP}}$ ða | [$_{\texttt{IP}}$ Crist | [mid drium fotwylmum] | [ofer hyre yþum] | [mihtelice] | *eode* ]] | |
| ... then | when | Christ | with dry foot-steps | over its waves | miraculously | went | |

'... when Christ went over its waves miraculously with dry footsteps'
(cocathom1,ÆCHom_I,_7:235.94.1258)

As before, I collected clauses with and without non-finite verbs separately. The X constituent in the *SXV* and *SVX* patterns was then simply replaced by the non-finite verb. I did not include illustrations of clauses with non-finite verbs above in order to save space. Hence, I conducted another 3 (word order patterns) * 2 (subject type) * 2 (main vs. non-finite verb) * 2 (period; Old and Middle English) = 24 search queries.[6]

**Results**

The rates of verb-final and verb-medial structures in SCs can now be used to adjust the counts of these word orders in MCs and CCs. I will first discuss the calculation for sentences with pronominal subjects. The data is presented in table 3.6.

| Clause Type | *V S* High verb `V-to-C` | *SVX* Verb-medial Potentially `I`-initial | *SXV* Verb-final `I`-final | Adjusted Potentially `I`-initial | Adjusted `I`-final |
|---|---|---|---|---|---|
| MC | 7,629 (41.2%) | 10,125 (54.7%) | 766 (4.1%) | 13,661 (73.9%) | 4,820 (26.1%) |
| CC | 1,528 (15.7%) | 6,484 (66.7%) | 1,716 (17.6%) | 7,192 (74.0%) | 2,528 (26.0%) |
| SC | 128 (0.5%) | 11,678 (46.4%) | 13,388 (53.1%) | | |

Table 3.6: Adjusted rates of verb-final and verb-medial MCs and CCs with pronominal subjects

The sampled SCs show high verb placement, verb-medial structures and verb-final patterns in 0.5%, 46.4% and 53.1% of all cases respectively. As expected, instances of high verb placement are so rare, 0.5%, that the measurement yields a virtually unimpaired estimate of the rates of verb-final and verb-medial clauses. I then distributed the MCs and CCs with verb placement under `C` over the two remaining word order types in accordance with the relative frequencies found in subordinate clauses. For example, 46.4% of all subordinate clauses with pronoun subjects are verb-medial. Hence, one might assume that if a main clause verb that occurs under `C` had not been placed there, it would have had a 46.4% chance of surfacing in a potentially `I`-initial clause. I therefore took 46.4% of the 7,629 = 3,536 high verb sentences in MCs and added them to the original 10,125 verb-final sentences found in that clause type. This yielded a total of 13,661 adjusted potentially `I`-initial clauses. The algorithm for this calculation is sketched with arrows in table 3.6 above. I then repeated the calculation for CCs and clauses with `I`-final configurations. As hypothesized, this adjustment results in a substantial convergence in the rates of verb-final and verb-medial patterns between MCs and CCs. In fact, the adjusted rates are now surprisingly similar. The correction suggests that about 26.1% of MCs and 26.0% of CCs are underlyinly `I`-final. The difference is not statistically significant ($\chi^2$=0.014, df=1, $p$=0.9063). This finding supports the hypothesis that it is at least plausible to assume identity in the underlying rate of `I`-final headedness in MCs and CCs.

[6]The CorpusSearch query and output files can be found on the Dissertation DVD under `Chapter 3 - Conjunct Clauses/03 Hypothesis Testing/H2`.

I now turn to the same measurement with full subjects. The relevant data is given in table 3.7.

| Clause Type | *V S* (structurally ambiguous) | *SVX* Verb-medial Potentially I-initial | *SXV* Verb-final I-final | Adjusted Potentially I-initial | Adjusted I-final |
|---|---|---|---|---|---|
| MC | 9,706 (55.6%) | 7,010 (40.1%) | 751 (4.3%) | 112,778 (74.5%) | 4,199 (25.5%) |
| CC | 2,358 (35.3%) | 3,496 (52.3%) | 825 (12.4%) | 4,776 (74.2%) | 1,663 (25.8%) |
| SC | 991 (10.2%) | 5,270 (54.3%) | 3,449 (35.5%) | | |

Table 3.7: Adjusted rates of verb-final and verb-medial MCs and CCs with full subjects

The sample of SCs with full subjects shows a proportion of 10.2%, 54.3% and 35.5% of V-to-C sentences, potentially I-initial structures and clauses with I-final alignment respectively. The relative frequency of *VS* clauses is noticeably higher in SCs with full subjects than those with pronominal subjects (10.2% vs. 0.5%). Again, this discrepancy can be attributed to a large number of post-verbal full subjects that are placed in the low subject position. Consequently, the measurement with full subjects will be somewhat less reliable than the previous one, but it should still allow for a relatively adequate estimation of the proportions of verb-final and verb-medial clauses. I took the SC rates thus obtained as a factor for the apportionment of high verb structures in MCs and CCs into the remaining two word order groups. The calculation was executed as in the previous context. The results show, once more, a considerable harmonization in the rates of I-final clauses and structures with other IP-headedness among MCs and CCs. The value of adjusted I-final clauses in MCs is 25.5%, in CCs 25.8%. The difference is not statistically significant ($\chi^2$=0.268, df=1, $p$=0.6044). It is therefore reasonable to believe that MCs and CCs may occur with underlyingly identical rates of I-initial and I-final phrase structure.

I was somewhat surprised by the high degree to which the relative frequency of verb-medial and verb-final patterns converged based on the proportions of these patterns in subordinate clauses. If the rates of I-final structures are indeed identical in MCs and CCs, this would mean that the hypothetical distribution of IP configurations that did not materialize because the verb was instead placed under C corresponds remarkably well to the rates of IP-configurations in embedded clauses.

If true, what would be the meaning of this? Is there a latent link between the rates of I-final headedness in main and subordinate clauses? If the rate of IP-headedness in SCs plays a role for V-to-C structures in MCs and CCs, why does it not also influence clauses in which the verb remains in I in those clauses? Might it be possible that the distribution of high verb structures over the remaining two patterns does not actually follow the respective rates in SCs, but occurs randomly with a probability of 50% for each? This allocation would yield very similar results. Unfortunately, I do not know the answers to these questions. I will therefore not contemplate possible implications of this finding further.

What is important is that if verbs that occur under C are distributed over I-medial and I-final patterns according to SC rates (53% vs. 47% with pronoun subjects) or randomly (50% - 50%), then MCs and CCs would have the same underlying rates of verb-final and verb-medial phrase structure, and the potential for this identity was all Hypothesis 2 asserted.

In summary, apportioning high verb placement structures over verb-medial and verb-final structures according to the respective rates found in subordinate clauses leads to a substantial convergence in the proportions of these word order patterns between MCs and CCs. I will therefore accept the second hypothesis as well. My model explains this observation as a consequence of actual identity in the underlying rates of verb-final and verb-medial structures in MCs and CCs. CCs show a greater number of verb-final patterns merely because they show the verb after the subject more frequently overall and not because they have an inherently higher rate of I-final headedness. The higher rate of verbs positioned inside the IP in CCs, in turn, can be explained as a consequence of C-head conjunctions blocking high verb placement in that clause type.

#### 3.4.3.3 Hypothesis 3: Differential Development of V-to-C Contexts

High verb placement under C in declarative matrix clauses gradually declines during the Early English period. My model predicts that this decline should be slower in CCs than MCs. This can be tested by regressing the proportion of 'verb - pronominal subject' patterns against time and testing for the significance of a time - clause type interaction effect.

**Data Collection**

For this study, I compared all clauses with a pronominal subject and a preceding finite verb to those with a following finite verb as surface realizations of V-to-C and verb-in-IP structures respectively. I separated clauses without a conjunction, MCs, from those that featured such an element before the subject, CCs. I did not collect any new data, but re-used the data set compiled for the first investigation, summarized in table 3.4 above. The first column of this table, high verbs, contains all instances of post-verbal pronominal subjects and are thus representative of clauses with verbs placed under C. The second and third columns, verb-medial and verb-final patterns, were grouped together and thus represent clauses with the verb inside the IP.

The following sentences illustrate the four resulting word order types that are relevant for this study.

(128) **MC**
**+ V-to-C:**
*V - Pronominal subject*

Ða [$_{\text{C}}$ *geseah* ] [$_{\text{IP}}$ he Marinan þæt mægden ]
then saw he Mary that maiden

'Then, he saw the girl Marina' (comart3,Mart_5_[Kotzor]:Jy7,B.11.1097)

(129) **CC**
**+ V-to-C:**
*Conjunction ... V - Pronominal subject*

& þa [$_{\text{C}}$ *geseah* ] [$_{\text{IP}}$ ic ðær þone rumestan feld ]
and then saw I there the most-spacious field

'And then I saw the most spacious field there' (cobede,Bede_5:13.430.2.4323)

(130) **MC**
**+ verb in IP:**
*Pronominal subject ... V*

[$_{\text{IP}}$ Ic hyne *geseah* on þam ylcan flæsce ]
I him saw in the same flesh

'I saw him in the same flesh' (covinsal,VSal_1_[Cross]:21.12.177)

(131) **CC**
**+ verb in IP:**
*Conjunction ... Pronominal subject ... V*

**and** [$_{\text{IP}}$ he ða *geseah* sumne dæl þæs leohtes ]
and he then saw some part of-the light

'And he then saw some part of the light' (cocathom2,ÆCHom_II,_11:107.528.2325)

All hits were then annotated for their source file and its approximate year of composition as recorded in tables 3.1, 3.2 and 3.3.[7]

[7] The Hypothesis 3 data is summarized as a wide format excel file, which can be found on the Dissertation DVD under `Chapter 3 - Conjunct Clauses/03 Hypothesis Testing/H3/H3 Summary.xlsx`.

**Results**

I will first demonstrate that high verb placement does indeed become less frequent during the course of the Old and early Middle English periods. Table 3.8 presents a simple logistic regression model that predicts the realization of the verb before pronominal subjects from the time variable alone.

```
formula = HighVerb ~  Year, family = binomial, data = H3

             Estimate      Std.Error z-value p
Intercept    2.9511        0.1742    17.04   <0.001***
Year         -0.0037651    0.0001772 -21.25  <0.001***

Null deviance:          35591 on 28247 degrees of freedom
Residual deviance:      35109 on 28246 degrees of freedom
AIC: 35113
```

Table 3.8: Logistic Regression Model 1: high verb placement against time

The model shows an unmistakable decline in verb placement under `C` during the Early English period. Specifically, it predicts that the log-odds of finding a `V`-to-`C` structure decline by 0.0038 per year. Almost half of all finite verbs are estimated to be placed in the high verb position in the $9^{\text{th}}$ century, but only about one in five verbs should occur there by the beginning of the $13^{\text{th}}$ century. The 'Year' coefficient is a highly significant predictor for verb positioning in this model.

Next, I shall test the hypothesis at hand. For this purpose, I will fit a mixed-effects model to the high verb placement data. It predicts the probability of finding a verb before a pronominal subject from time, clause type (MCs vs. CCs) as well as their interaction as fixed effects, and includes varying intercepts for every text to control for random text variability. The result is shown in table 3.9.

```
formula = HighVerb ~  Year + ClauseType + Year:ClauseType + (1 | Text),
              family = binomial, data = H3

Fixed effects:
                         Estimate      Std.Error     z-value p
Intercept                4.8742        0.8100        6.018   <0.001***
Year                     -0.0052681    0.0008065     -6.532  <0.001***
ClauseType(MC→CC)        -4.3235       0.4304        -10.046 <0.001***
Year:ClauseType(MC→CC)   0.0028894     0.0004392     6.579   <0.001***


Random effect:
Text, N=101
Variance of random intercepts:  0.785

Null deviance:          35591 on 28247 degrees of freedom
Residual deviance:      31082 on 28243 degrees of freedom
AIC: 31092
```

Table 3.9: Logistic Regression Model 2: mixed-effects model for high verb placement by time and clause type

The model predicts a decrease in the log-odds of high verb placement by -0.0053 per year in MCs. This confirms, once again, the gradual decline of `C` as a verb position in declaratives during the Early English period. As the clause type category is changed from MCs to CCs, the overall log-odds of finding a `V`-to-`C` structure fall by about -4.3. This is in accordance with the earlier finding showing that the relative frequency of high verbs is significantly lower in CCs than in MCs (see table 3.4). Both of these predictors are highly significant in the model. Now, the most crucial aspect of the model concerns the interaction between the time and clause type variables. The model estimates that the loss of `V`-to-`C` structures is decelerated by 0.0029 log-odds per year in CCs by comparison with the rate of change in MCs. Thus, MCs lose high verb placement faster, at a rate of -0.0053, than CCs, at a rate of only -0.0053 +0.0029 = -0.0024. This effect conforms precisely to my model expectations according to which the decline in high verb placement should proceed more slowly in CCs than MCs. The interaction term, too, is a highly significant predictor for verb placement in this model. A likelihood ratio test on the difference in deviance between a reduced model without, and the full model with, the time -

clause type interaction term offers good evidence that the interaction term improves the model fit sufficiently to necessitate its inclusion in the model ($\chi^2$=39.44, df=1, $p$<0.001). I interpret these results to show that there is good evidence in favor of the proposed syntactic explanation of the difference between MCs and CCs. The high verb placement data comes from 101 text files overall. Their random intercepts show a variance of 0.785, which indicates to me moderate variability among texts.

As a side note, I found that failure to control for random variability among texts in the present study would likely overestimate the fixed effects. In particular, the time - clause type interaction term would become so strong that it would actually show a rise in V-to-C-structures in CCs. This would suggest that the loss of C-head conjunctions proceeds much more rapidly than the decline of high verb placement so that the former change overcompensates for the latter. In contrast, the mixed-effects model above suggest a decline in 'verb - pronominal subject' declaratives in both MCs and CCs, but a steeper slope in the first and a more restrained development in the second clause type. This differential seems to me to constitute a more judicious interpretation of the data.

All in all, the model performance is mediocre. It fits significantly better than a null model (where the null model is an intercept only model) (Likelihood Ratio Test, $\chi^2$=4869.7, df=4, $p$<0.001). However, it fits the data only approximatively (Pseudo-$R^2_{marginal}$= 0.134, Pseudo-$R^2_{conditional}$=0.301). The model has little classificatory power over the baseline (C-index=0.748, classification accuracy =72.9% vs. baseline: 67.6%). The relative poorness of this model is typical of models representing linguistic changes from the middle ages. The data from medieval languages are generally of a low quality, containing noise from sources such as translation effects, complex manuscript transmission histories, uncertain dating methods, inadequate corpus annotations, formulaic or stilted styles, and several other factors. Following the classical dictum that historical linguistics must make the best use of bad data, I shall accept the model despite its deficiencies and warn against over-interpreting its results.

The graph in figure 3.4 illustrates the model. It depicts the proportion of high verb placement with a green data point for MCs and a red data point for CCs in each of the 101 corpus files included in this study. On average, the red CC points fall below their green MCs counterparts, which reflects their lower overall probability of showing a finite verb before a pronominal subject. The green and red lines represent the model estimate for the development of high verb placement in MCs and CCs respectively as well as their 95%-confidence intervals. The green line for MCs runs more steeply than the red line for CCs, which indicates the significant difference in their slope. By c. 1200, the overall probabilities of high verb placement in MCs and CCs have almost completely converged.

To sum up, I could produce evidence for the predicted differential loss of V-to-C structures: CCs do indeed seem to lose high verb placement somewhat more slowly than MCs. I will therefore accept Hypothesis 3 as well. My model can explain this finding as an interaction effect between the loss of C-head conjunctions and the decline in high verb placement. The former development offsets the surface effects of the latter change to some degree.

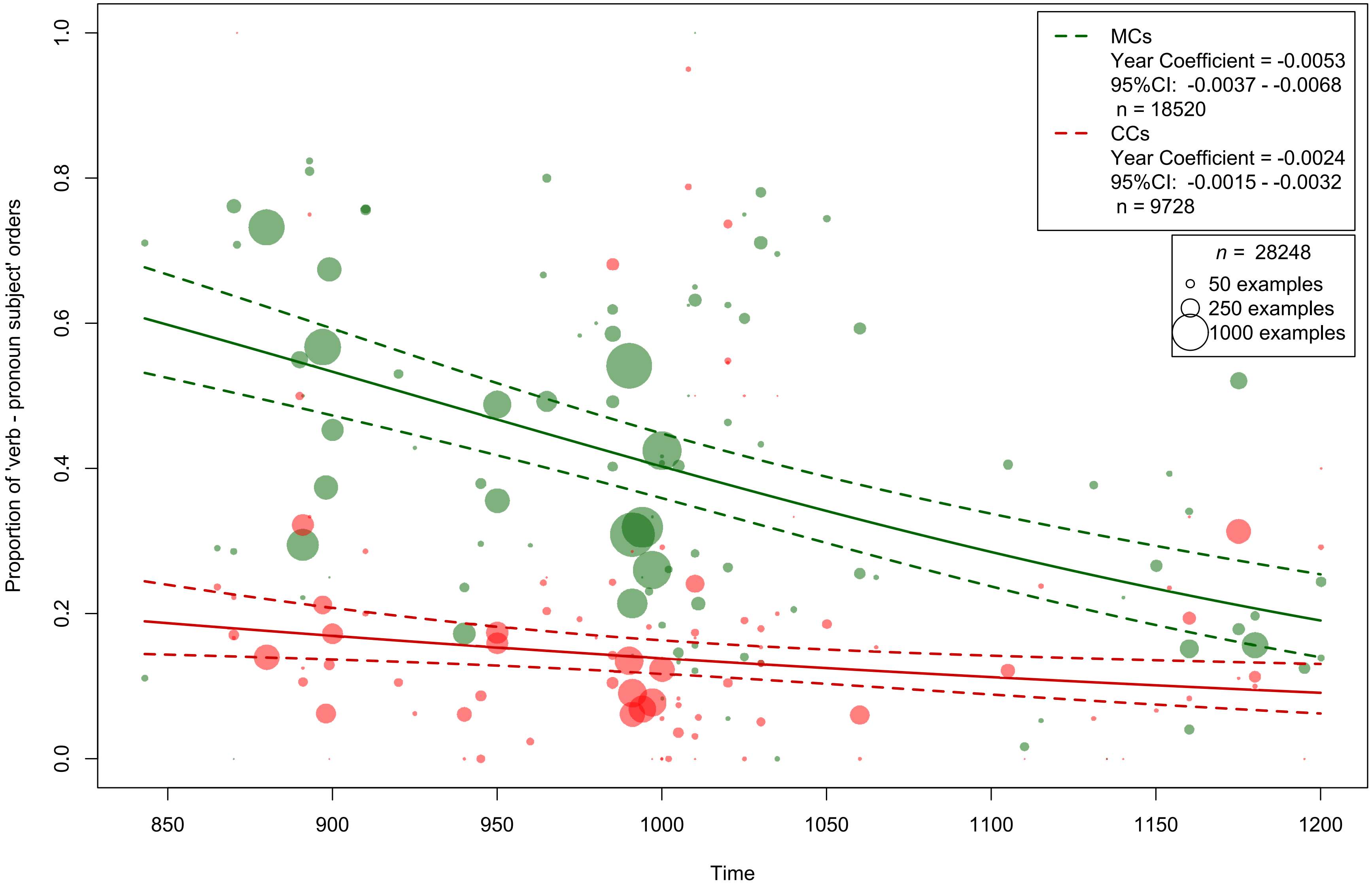



Figure 3.4: Illustration of the mixed-effects model for the decline in high verb placement

#### 3.4.3.4 Hypothesis 4: Effect of Separation of Conjunction and IP

According to my model, the elevated rate of verb-final structures in CCs should not be due to all conjunctions in general but specifically to those that can be parsed as C-heads. This hypothesis can be tested by separating CCs into cases with unambiguous logical connectors and cases that potentially feature C-head conjunctions. Subsequently, one can compare the relative frequency of verb-final patterns in those subtypes.

**Data Collection**

I subdivided the CC occurrences reported in tables 3.4 and 3.5 above into those clauses that necessarily involve logical connectors vs. those clauses that may potentially include C-head conjunctions. CCs were divided for all three recorded word orders, i.e., for the counts of high verb, verb-medial and verb-final CCs.

I will first discuss the partitioning of CCs with subject pronouns, i.e., the second row in table 3.4. High verb placement in CCs is necessitated by the order 'verb - subject pronoun' because pronominal subjects must be placed in the specifier of IP so that a preceding finite verb must occur under C. A conjunction in front of a verb under C must therefore be a logical connector, [conjunction ... *verb* [$_{IP}$ pronoun subject ... ]], (132). A conjunction before a high verb cannot possibly function as a C-head conjunction, (133).

(132) **Necessarily logical connector**
**+ High verb placement / V-to-C (with subject pronoun):**
*Conjunction ... Verb - Pronominal subject*

**Ond** þa æfter feawa dagum [$_{CP}$ ða [$_{C}$ *endode* ] [$_{IP}$ he his lif þurh martyrhad for Criste ]]
and then after few days then ended he his life through martyrdom for Christ

'And then, after a few days, he ended his life through a martyrdom for Christ'
(comart3,Mart_5_[Kotzor]:Se23,A.7.1809)

(133) **Potential C-head conjunction**
**+ High verb placement / V-to-C (with subject pronoun):**
{*Conjunction AND Verb* } *- Pronominal subject*
*impossible*

Verb-medial CCs must be parsed with logical connectors if they are separated from subject pronouns by another constituent, (134). They may involve C-head conjunctions if the conjunction is immediately adjacent to the pronominal subject, (135a). Furthermore, my model also allows for a C-head conjunction parse if only non-subject pronouns intervene between conjunction and the indefinite subject pronoun *man*. The reason is that non-subject pronouns can prepose to the CP-IP boundary by rule (81) and still leave C as a position for the conjunction available. I therefore allowed pronouns to intervene between conjunction and subject, as exemplified in (135b).

(134) **Necessarily logical connector**
**+ Verb-medial / Potentially I-initial:**
*Conjunction ... Some Constituent ... Pronominal subject - Verb*

**Ond** [$_{PP}$ on ðæm ðeowdome ] [$_{IP}$ he *geendade* his lif ]
and in the service he ended his life

'And in that service [=as a zoo keeper], he ended his life' (comart3,Mart_5_[Kotzor]:Ja16,A.9.103)

(135) **Potential C-head conjunction**
**+ Verb-medial / Potentially I-initial:**

a. *Conjunction - Pronominal subject - Verb ... X*

**And** [$_{IP}$ he swa *geendode* þa ealdan gecyðnysse ]
and he so ended the old testament

'And thus he ended the Old Testament' (colwstan1,ÆLet_2_[Wulfstan_1]:127.179)

b. *Conjunction - Pronoun - Pronominal subject - Verb ... X*

**&** [$_{IP}$ him [$_{IP}$ man *selle* an half swulung an Ciollandene ]]
and him one give a half sulung in Chillenden

'And one should give him half a sulung [= c. 60 acres of land] at Chillenden [=village in East Kent]'
(codocu1,Ch_1482_[HarmD_2:10.25)

The same is true for verb-final CCs. Where conjunction and subject pronoun are separated from each other, the clause must be parsed with logical coordination, (136), whereas adjacency of conjunction and pronominal subject opens up the possibility for the presence of a C-head conjunction, (137). Again, other pronouns were allowed to intervene between conjunction and subject for the second option.

(136) **Necessarily logical connector**
**+ Verb-final / Necessarily `I`-final:**
*Conjunction ... Some constituent ... Pronominal subject ... X ... Verb*

**ac** [$_{\text{PP}}$ þurh oþerne martyrdom] [$_{\text{IP}}$ hi heora lif *geendedon* ]
but through other martyrdom they their life ended
'But through another martyrdom, they ended their life' (comart3,Mart_5_[Kotzor]:Jy30,A.9.1315)

(137) **Potential C-head conjunction**
**+ Verb-final / Necessarily `I`-final:**
*Conjunction - Pronominal subject ... X ... Verb*

**and** [$_{\text{IP}}$ he ða mid geleafan his lif *geendode* ]
and he then with belief his life ended
'And he then ended his life with faith' (coaelive,ÆLS_[Maccabees]:104.4880)

I will now turn to CCs with full subjects, i.e., the second row in table 3.5. CCs with finite verbs in front of full subjects are usually only compatible with logical connectors. Either they involve verb placement under `C` in operator contexts, (138a), or an additional pre-verbal constituent, (138b). Both scenarios will render `C` unavailable for the conjunction.

(138) **Necessarily logical connector**
**+ High verb placement / `V-to-C` (with full subjects):**
*Conjunction ... Verb ... Full subject*

a. **&** þonne [$_{\text{C}}$ *wendeð* ] [$_{\text{IP}}$ him Sanctus Petrus þanon fram þære helle dura ]
and then turns reflexive Saint Peter thence from of-the hell door
'And then Saint Peter turns away from the door of hell'
(coverhom,HomU_6_[ScraggVerc_15]:193.2009)

b. **and** [$_{\text{PP}}$ æfter ðysum wordum ] [$_{\text{IP}}$ *gewende* se engel up ]
and after these words turned the angel up
'And after these words, the angel rose up (coaelive,ÆLS_[Exalt_of_Cross]:90.5623)

However, there is a context that quite plausibly allows for the presence of a C-head conjunction with full subjects. For this scenario, the finite verb must be positive and indicative so that it probably occurs inside the `IP`. For the Middle English data, I simply required the verb to be positive since mood distinctions are not annotated in the Middle English files. The full subject can now remain in the lower subject position and a conjunction immediately preceding the verb can potentially be positioned under `C`. An example of such a structure is given in (139a). Once again, I allowed pronouns to intervene between conjunction and finite verb, as illustrated in (139b).

(139) **Potential C-head conjunction**
**+ High verb placement / `V-to-C` (with full subjects):**

a. *Conjunction - Positive indicative verb - Full subject*

**&** [$_{\text{IP}}$ *wende* [$_{\text{VP}}$ Æðelstan hine eft into Sunnanbyrg ungebetra þinga ]]
and turned Athelstan reflexive again to Sunbury of-unimproved things
'And Athelstan returned again to Sunbury, things not having improved'
(codocu3,Ch_1447_[Rob_44]:17.78)

b. *Conjunction - Pronoun - Positive indicative verb - Full subject*

**ond** [$_{\text{IP}}$ him *toburston* [$_{\text{VP}}$ ða eagan ]]
and him off-burst the eyes
'And his eyes burst out off him' (comart3,Mart_5_[Kotzor]:Se8,B.16.1689)

Verb-medial and verb-final sentences with full subjects were treated exactly as those with pronominal subjects. Potentially `I`-initial CCs must involve a logical connector if the conjunction is separated from the subject by a non-pronominal constituent, (140), but may feature a C-head conjunction if the conjunction is immediately adjacent to the subject, (141a), or separated from it only by a pronoun, (141b). The examples below illustrate for clauses where the heavy constituent X was instantiated as a non-finite verb or infinitival clause.

(140) **Necessarily logical connector**
**+ Verb-medial / Potentially `I`-initial:**
*Conjunction ... Some Constituent ... Full subject -* $V_{fin}$ ... $V_{non-fin}$

& [$_{DP}$ ealle þa men ] [$_{IP}$ Iulius *het* ofslean ]
and all the men Julius ordered off-slay

'And Julius commanded that all those men be slain' (coorosiu,Or_5:12.128.6.2704)

(141) **Potential C-head conjunction**
**+ Verb-medial / Potentially `I`-initial:**

a. *Conjunction - Full subject -* $V_{fin}$ ... $V_{non-fin}$

& [$_{IP}$ se Cenwalh *het* atimbran þa ciricean on Wintunceastre ]
and the Cenwalh ordered build the church in Winchester

'And Cenwalh ordered the church in Winchester to be built'
(cochronA-1,ChronA_[Plummer]:643.1.322)

b. *Conjunction - Pronoun - Full subject -* $V_{fin}$ ... $V_{non-fin}$

& [$_{IP}$ him þæt heafod *wæs* adune gewended ]
and him that head was down turned

'And his head was turned downward' (coblick,LS_32_[PeterandPaul[BlHom_15]]:173.32.2178)

Necessarily `I`-final CCs feature a logical connector, (142), or a potential C-head conjunction, (143), under the same conditions. Again, the following illustrations involve clauses with non-finite verbs.

(142) **Necessarily logical connector**
**+ Verb-final / Necessarily `I`-final:**
*Conjunction ... Some constituent ... Full subject ...* $V_{non-fin}$ ... $V_{fin}$

& [$_{ADVP}$ eft ] [$_{PP}$ binnen seofen gearen ] [$_{IP}$ his ban up genumen *wurðen* ]
and again within seven years his bones up taken were

'and after seven years, his bones were taken up again' (coneot,LS_28_[Neot]:114.105)

(143) **Potential C-head conjunction**
**+ Verb-final / Necessarily `I`-final:**
*Conjunction - Full subject ...* $V_{non-fin}$ ... $V_{fin}$

& [$_{IP}$ se cyning heo ofslean *het* ]
and the king them off-slay ordered

'And the king ordered that they be slain' (cobede,Bede_4:18.308.13.3120)

I ran search queries that found conjunctions immediately adjacent to subjects, hence potential C-head conjunctions, on the output files for CCs from the investigation of the first hypothesis. The complement files of these searches contain all examples of conjunctions that are separated from the subject, hence unambiguous logical connectors. I ran a total of 3 (word order patterns) * 2 (subject type) * 2 (main vs. non-finite verb) * 2 (period; Old and Middle English) = 24 search queries.[8]

[8]The CorpusSearch query, output and complement files can be found on the Dissertation DVD under `Chapter 3 - Conjunct Clauses/03 Hypothesis Testing/H4`.

**Results**

Table 3.10 below presents the results of this study. It shows the relative frequencies of clauses with high verb placement under C, potentially I-initial configurations and necessarily I-final structures in MCs, CCs and, for comparative purposes, SCs. The table include clauses with pronominal and full subjects, i.e., it is essentially the unification of the counts from tables 3.4 and 3.5. The CC row is then divided into two, one for logical connectors, the other for potential C-head conjunctions, according to the criteria outlined above. These rows represent the new information critical for the evaluation of the present hypothesis.

| Clause type | | *V S* High verb (V-to-C) | *SVX* Verb-medial Potentially I-initial | *SXV* Verb-final I-final |
|---|---|---|---|---|
| MC | | 17,335 (48.2%) | 17,135 (47.6%) | **1,517 (4.2%)** |
| CC | | 3,886 (23.7%) | 9,980 (60.8%) | 2,541 (15.5%) |
| | necessarily logical connector | 3,486 (56.9%) | 2,283 (37.3%) | **353 (5.8%)** |
| | potentially C-head conjunction | 400 (3.9%) | 7,697 (74.8%) | **2,188 (21.3%)** |
| SC | | 1,119 (3.2%) | 16,948 (48.6%) | 16,837 (48.2%) |

Table 3.10: Verb positions in MCs, separated CCs, adjacent CCs and SCs, pronominal and full subjects

As expected, I-final clauses are significantly more common in clauses with potential C-head conjunctions, 21.3%, than in MCs, 4.2% ($\chi^2$=3157.6, df=1, $p$<0.001, odds ratio = 6.14, 95%-confidence interval [5.72 – 6.59]), or in clauses that necessarily involve a logical connector, 5.8% ($\chi^2$=703.9, df=1, $p$<0.001, odds ratio = 4.42, 95%-confidence interval [3.92 – 4.98]). This finding provides further evidence for the proposed analysis, in which the high rate of I-final CCs predominantly results from conjunctions that block C as a verb position and not from the presence of conjunctions generally.

There is also a significant difference in the rate of I-final headedness between CCs with logical connectors, 5.8%, and MCs, 4.2% ($\chi^2$=29.2, df=1, $p$<0.001). The statistical significance may be due to the large sample size. The clause type effect is not very large - the odds of finding an unambiguous logical connector with verb-final structure are only c. 1.4 times greater than those of finding MCs with this word order (odds ratio = 1.39, 95%-confidence interval [1.23 – 1.57]). Nevertheless, this finding is unexpected under my model, which would have predicted that if the conjunction must be a logical connector, the frequency of I-final structures should not differ substantially between CCs and MCs. Perhaps this discrepancy is caused by measuring error, by corpus annotation mistakes, by noise from parenthetical or disfluent interpretations of the sentence elements intervening between conjunction and IP or by available parses not accounted for by my model.

To summarize, this investigation supports the hypothesis that separating the conjunction from IP should have a strong effect on verb placement patterns. It is indeed the case that the finite verb appears in a configuration that indicates I-final headedness less frequently if the conjunction is separated from IP than when it is immediately adjacent to it. This supports the analysis of conjunctions in front of IP as potential C-heads. There remains a significant difference in I-final headedness between MCs and CCs necessarily including logical connectors. More work is required to uncover the cause of this unexplained divergence. Overall, the data is in accordance with the predicted word order distributions and one is thus justified in accepting Hypothesis 4.

#### 3.4.3.5 Hypothesis 5: Differential Development of I-Final Structures

Verb-final structures are lost in English during the course of the Old and early Middle English periods. My grammar model predicts that this development should proceed somewhat faster in clauses that may involve a C-head conjunction than in all other clause types. This hypothesis can be tested by regressing the distribution of *SXV* vs. *VS* plus *SVX* patterns in various clause types against time.

**Data Collection**

I did not collect any new data for the investigation of this hypothesis. Instead, I re-used the data collected for the preceding hypothesis shown in table 3.10. The *SXV* column includes the counts of IPs that must necessarily be head-final. I combined the *VS* and *SVX* columns to create the counter-context, i.e., IPs that do not necessarily show a head-final configuration. The 128 instances of high verb placement in subordinate clauses with pronominal subjects (see table 3.6) are very likely to involve root clause syntax (see ex. (124a)). I therefore allocated these examples to the *VS* column for MCs. Every example was then annotated for the text file it was found in and its approximate year of composition according to tables 3.1, 3.2 and 3.3.[9]

**Results**

I fitted a mixed-effects logistic regression to the data set for I-final clauses. It predicts the occurrence of a verb-final clause (vs. a high verb / verb-medial clause) from its assigned date, clause type (MCs, CCs with logical connectors, abbreviated CC log. con., CCs with potential C-head conjunctions, abbreviated CC C-head c., and SCs), and the interaction between these two variables, while controlling for random variation among texts. The resulting model is presented in table 3.11.

```
formula = VerbFinal ~  Year + ClauseType + Year:ClauseType + (1 | Text),
             family = binomial, data = H5

Fixed effects:
                                    Estimate     Std.Error z-value p
Intercept                           2.9167       1.1904    2.450   0.01428*
Year                                -0.0063455   0.0011922 -5.323  <0.001***
ClauseType(MC→CC(log.  con.))       0.8945       1.2800    0.699   0.48466
ClauseType(MC→CC(C-head c.))        3.1561       0.4743    6.655   <0.001***
ClauseType(MC→SC)                   3.1017       0.4398    7.058   <0.001***
Year:ClauseType(MC→CC(log.  con.)) -0.0006035    0.0013266 -0.455  0.64914
Year:ClauseType(MC→CC(C-head c.))  -0.0012987    0.0004896 -2.652  0.00799**
Year:ClauseType(MC→SC)              0.0001689    0.0004550 0.371   0.71045

Random effect:
Text, N=101
Variance of random intercepts:  0.6619

Null deviance:          96086 on 87297 degrees of freedom
Residual deviance:      67757 on 87289 degrees of freedom
AIC: 67775
```

Table 3.11: Logistic Regression Model 3: mixed-effects model for verb-final structures by time and clause type

The model estimates that MCs lose necessarily I-final structures at a rate of -0.0063 log-odds of *SXV* per year. The time variable is highly significant in this model, which is well in line with the expectation that verb-final structures should be going extinct during the period under investigation here. As the clause type category is changed from MCs to CCs with logical connectors, the overall log-odds of finding an I-final clause increase ever so slightly by 0.89. This finding agrees with both the empirical observation above that the proportions of this word order are comparable between MCs and CCs with logical connectors as well as with my theoretical conjecture that they should in fact be identical in both clause types. The difference is not statistically significant in this model. Changing the clause type from MCs to CCs with potential C-head conjunctions or to SCs raises the probability of *SXV* clauses by log-odds of 3.16 and 3.10 respectively. This result reflects the fact that verb-final patterns are much less likely in MCs than in the latter two clause types. Both clause type differences are

[9] The Hypothesis 5 data is summarized as a wide format excel file, stored on the Dissertation DVD under `Chapter 3 - Conjunct Clauses/03 Hypothesis Testing/H5/H5 Summary.xlsx`.

highly significant in this model. The subsequent three variables represent the adjustment in the rate of change as the clause type variable is changed from MC to some other clause type. They are the crucial effects for the investigation of this hypothesis. The loss of `I`-final orders does not proceed significantly more slowly or rapidly in CCs with logical connectors ($p$=0.649). Nor does it progress at a significantly different rate in SCs ($p$=0.710). However, the model predicts that the difference in the rate of the loss of *SXV* patterns is in fact significant at the 1%-level for CCs with potential C-head conjunctions ($p$=0.008). Specifically, while the log-odds of finding a verb-final clause decline by only -0.0063 per year in MCs, they do so at a faster rate of -0.0063 + -0.0013 = -0.0076 in CCs with potential C-head conjunctions. A likelihood ratio test on the difference in deviance between a reduced model without, and the full model with, the time - clause type interaction terms shows that the inclusion of the interaction terms improves the model sufficiently to justify their inclusion ($\chi^2$=10.24, df=3, $p$=0.0166). These results corresponds exactly to my expectations. As predicted, there is evidence to suggest that CCs with potential C-head conjunctions lose `I`-final structures faster than MCs, whereas MCs, CCs with logical connectors and SCs all undergo the development at the same rate in accordance with the Constant Rate Hypothesis.

It is worth mentioning, once more, that the present study requires some control for random variability among texts. If random intercepts for every text were not included in the regression model, the effects of some variables would be over or under-estimated. For example, the overall proportion of *SXV* clauses in SCs remains relatively stable during the Anglo-Saxon era, but shows a dramatic drop when English texts slowly re-emerge after the document hiatus caused by the Norman Conquest. A fixed-effects model would therefore steepen the slope in SCs relative to a model that can adjust its slope in accordance with the variability among texts found during a particular period. More importantly, the omission of varying text intercepts would cause the slope difference between MCs and potential C-head CCs to miss the 5% significance level, albeit only barely.

The model performs in an acceptable way. It fits significantly better than a null model (where the null model is an intercept only model) (Likelihood Ratio Test, $\chi^2$=28746.3, df=8, $p$<0.001). The model fit is relatively good given that it is based on patchy data from the middle ages (Pseudo-$R^2_{marginal}$= 0.396, Pseudo-$R^2_{conditional}$=0.498). The model has good classificatory power but improves little over the baseline (C-index= 0.854, classification accuracy= 80.3% vs. baseline: 76.0%).

The model is illustrated in the graph in figure 3.5. It shows the proportion of verb-final clauses in the 101 corpus text files separately for each of the four clause types, the model regression lines and their 95%-confidence intervals. MCs and CCs with logical connectors are shown in dark and light green, CCs with potential C-head conjunctions in red and SCs in blue. The first two clause types show a low probability of encountering a verb-final structure and their regression lines are largely overlapping. This shows that these clauses do not behave significantly different from one another. The red potential C-head and blue SC regression lines run above the green lines. Therefore, these clause types have a substantially higher probability of realizing a verb-final structure than MCs. Finally, one can discern a steeper slope for the red line than for the other three. This underlines the fact that the loss of *SXV* patterns proceeds significantly faster in CCs with potential C-head conjunctions than in the other clause types. Verb-final main clauses have virtually become extinct by 1200, so that there remains no more evidence for the differential behavior between MCs and CCs regarding this word order pattern from that time on.

Summing up, the regression model in table 3.11 supports the fifth hypothesis. It shows that CCs with potential C-head conjunctions do indeed seem to lose verb-final clauses faster than all the other clause types. This finding may be the most imposing of the results presented in this part. It conforms to a very risky prediction that could easily not have come true. One should of course remain cautious about the outcome of this investigation, for instance because of uncertainties in text dating, the potential for measuring error, possible influences from other factors such as subject type, genre, verb type etc., the effect of the individual conjunction lexeme, alternative statistical modeling techniques, etc. Nevertheless, my grammar model, according to which the loss of C-head conjunctions should free up `C` as a position for finite verbs thus amplifying the disappearance of verbs surfacing under head-final `IP`s, should be considered much more seriously as the correct explanation for the special status of Old English conjunct clauses now that the result of this study has proven positive than if it had been negative.

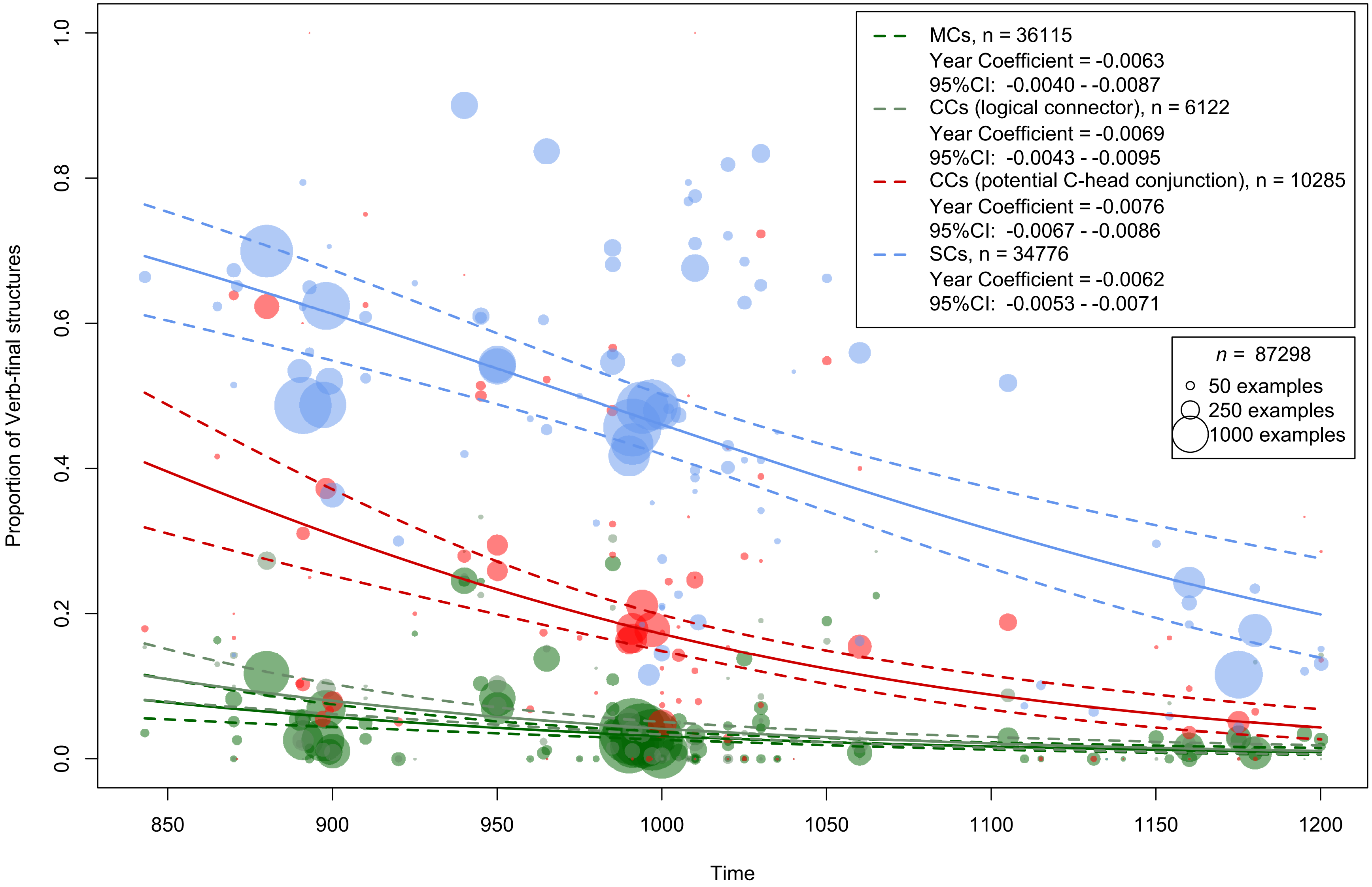



Figure 3.5: Illustration of the mixed-effects model for the decline in I-final structures

#### 3.4.3.6 Hypothesis 6: Uniform Development of Necessarily I-Initial Structures

The increase in I-initial structures as measured with post-verbal diagnostic elements should proceed at the same speed in MCs and CCs. The hypothesis is testable in a straightforward way - one must simply collect all MCs and CCs with diagnostic elements and regress the proportion of these elements in post-verbal position against time and clause time.

**Data Collection**

I collected all MCs and CCs with any overt subject, a subsequent finite verb, excluding non-finite verbs, and a diagnostic element. Diagnostic elements were pronouns, particles, stranded prepositions, or negatively quantified objects. I recorded whether the diagnostic occurred in pre- or post-verbal position. In the subsequent paragraphs, I will summarize the behavior of each of these elements with respect to their diagnostic value in post- and pre-verbal position based on findings reported in Pintzuk and Haeberli (2008), and mention other noteworthy points that emerged during the data collection process.

(i) **Pronouns**. Non-subject pronouns (possibly reinforced by *self*) are by far the most common diagnostic. Pronouns cannot postpose and thus necessarily indicate I-initial headedness in post-verbal position, (144).

(144) **Necessarily I-initial**
**Diagnostic: Pronouns**

a. *Subject ... V ... Pronoun*

þu [$_{I'}$ *tehtest* [$_{VP}$ **me** eac swa ungelygena gewittnesse ]]
you showed me also so un-lied witness

'You have also shown me such a reliable testimony' (cosolilo,Solil_3:66.26.928)

b. *Conjunction ... Subject ... V ... Pronoun*

**and** ðu [$_{I'}$ *mærsast* [$_{VP}$ **me** ]]
and you glorify me

'And you glorify me' (cocathom2,ÆCHom_II,_9:76.147.1547)

Pronouns frequently prepose. Therefore, pre-verbal pronouns often remain compatible with an I-initial parse in pre-verbal position. Such an I-initial parse is sketched with labeled brackets in (145a)-(145b). They only force I-final phrase structure if they co-occur with other pre-verbal material, (145c)-(145d).

(145) **Not necessarily I-initial**
**Diagnostic: Pronouns**

a. *Subject ... Pronoun ... V*

þu **me** [$_{I'}$ *ahsast* [$_{VP}$ micles & earfoðes to ongitanne ]]
you me ask great and difficult to understand

'You ask me a great and difficult [thing] to understand' (coboeth,Bo:42.147.23.2943)

b. *Conjunction ... Subject ... Pronoun ... V*

**&** þeah þu **me** [$_{I'}$ *tæhtest* [$_{VP}$ ær þa duru ]]
and though you me showed earlier the door

'and yet you showed me the door earlier' (coboeth,Bo:35.97.16.1878)

c. *Subject ... X ... Pronoun ... V*

we [$_{I'}$ [$_{VP}$ nu on idlum gilpe **us** mid golde & mid gimmum ] *gearwiaþ* ]]
we now in idle pride us with gold and with gems prepare

'We now adorn ourselves with gold and with gems in idle pride'
(coverhom,HomU_8_[ScraggVerc_2]:98.372)

d. *Conjunction ... Subject ... X ... Pronoun ... V*

**Ac** we [$_{I'}$ [$_{VP}$ þa mid scyldum **us** ] *scyldan* ]]
But we then with shields us shield

'But we would then shield ourselves with our shields' (coalex,Alex:17.10.190)

Hence, all post-verbal pronouns necessarily indicate I-initial headedness, but the diagnostic value of pre-verbal pronouns for I-final headedness is limited.

(ii) **Particles**. Adverbial particles do not postpose and can thus function as diagnostics for the present study. They necessarily indicate `I`-initial phrase structure post-verbally, as illustrated with the particle *forth* in (146a)-(146b). The Middle English text files annotate several items that are treated as particles in Old English as adverbs heading directional or locative adverb phrases. Examples of such adverbs are *forth* or *away*. I therefore queried directional and locative adverb phrases for the Middle English data as well. The Middle English sentences in (146c) - (146d) exemplify, once more, the particle (or adverb) *forth* in post-verbal position.

(146) **Necessarily `I`-initial**
**Diagnostic: Particles**

a. *Subject ... V ... Particle*

Sumes weliges mannes æcer [$_{I'}$ *brohte* [$_{VP}$ **forð** gode wæstmas ]]
some rich man's field brought forth good fruits

'The field of a rich man brought forth a good harvest' (cowsgosp,Lk_[WSCp]:12.16.4665)

b. *Conjunction ... Subject ... V ... Particle*

**&** Cynric his sunu [$_{I'}$ *ricsode* [$_{VP}$ **forþ** xxvi wintra ]]
and Cynric his son ruled forth 26 winters

'And Cynric, his son, reigned for twenty-six years' (cochronA-1,ChronA_[Plummer]:534.1.185)

c. *Subject ... V ... Particle*

þe hali gast, he [$_{I'}$ cumþ [$_{VP}$ **forþ** of hem bam ]]
the holy ghost he comes forth of them both

'As for the Holy Ghost, he comes forth from both of them' (CMVICES1,25.277)

d. *Conjunction ... Subject ... V ... Particle*

and al þos world [$_{I'}$ wende [$_{VP}$ **forþ** ]]
and all this world turned forth

'And this whole world [=the era of the Old Testament] went by' (CMLAMB1,81.146)

Clauses with pre-verbal, specifically prefixed, particles are compatible with an `I`-initial analysis. Example (147) shows the particle verb *of-slogon* 'slew off' before a pronoun, *hine*, which forces `I`-initial alignment.

(147) Hi þa [$_{I'}$ **of**-*slogon* [$_{VP}$ **hine** ]]
They then off-slew him

'Then, they killed him' (cowsgosp,Mk_[WSCp]:12.8.3113)

I therefore decided to collect particles only if they were spelled separately, hence annotated as independent clause functions, `RP`|`RPX`, and not spelled as one word along with the finite verb, and hence coded as prefixes to a verbal stem, e.g., `RP+VBPI`|`RP+VBDI` etc. I hope that in this way, I could minimize the chances of particle placement outside of the `VP`. Furthermore, unlike the YCOE, the Middle English data does not annotate particle prefixes on verbs. For instance, the Middle English corpus annotates *of-slagen* 'slain off' (CMPETERB1,48.199) simply as `VAN`, and not as `RP+VAN`. Hence restricting the search queries to orthographically distinct particles will also ensure comparability across the Old and early Middle English periods. Examples of pre-verbal free-form particles are illustrated in (148), again with the item *forth*.

(148) **Not necessarily `I`-initial**
**Diagnostic: Particles**

a. *Subject ... Particle ... V*

Eac swylce heo [$_{I'}$ [$_{VP}$ ða monoðlican **forþ** ] *gecigeþ* ]
also such it the menstruation forth invokes

'Likewise, it [=the plant *uiola*] brings on menstruation' (coherbar,Lch_I_[Herb]:165.1.2415)

b. *Conjunction ... Subject ... Particle ... V*

**&** he [$_{I'}$ [$_{VP}$ monige **forð** ] *acigde* ]
and he many forth summoned

'And he incited many [to repent for their sins]' (cobede,Bede_5:15.444.23.4472)

Hence, orthographically distinct particles force `I`-initial headedness in post-verbal position. They have at least a high, albeit not perfect, probability of occurring inside the `VP` and thus additionally diagnose `I`-final phrase structure in pre-verbal position.

(iii) **Stranded prepositions**. Stranded prepositions cannot postpose and thus indicate `I`-initial headedness in post-verbal position, (149). Prepositions can be stranded in Early English main clauses, rarely, as a consequence of preposing the adverbial component of R-pronouns, (149a), and, much more commonly, because of the fronting of pronominal complements of prepositions, (149b). In the examples below, I informally sketch the dependencies between the preposition and its preposed complement with an arrow.

(149) **Necessarily `I`-initial**
**Diagnostic: Stranded prepositions**

a. *Subject ... V ... Stranded preposition*

þæs cræft ic þær [$_{I'}$ *herige* [$_{VP}$ **on** _ ]]

of-that-one skill I there praise in

'The skill of him [=the maker of beautiful clothes], I praise therein [= in this speech?]' (coboeth,Bo:14.30.22.536)

b. *Conjunction ... Subject ... V ... Stranded preposition*

**and** Beorn eorl him [$_{I'}$ *for* [$_{VP}$ **mid** _ ]]

and Beorn earl him went with

'And Beorn the earl went with him' (ChronD_[Classen-Harm]:1050.16.1869)

Stranded prepositions are more inert than pronouns or particles. They are therefore very likely to indicate the position of `VP` in general and `I`-final headedness when positioned pre-verbally in particular. Relevant examples are shown in (150). The labeled bracketing structures indicate just one of several possible analyses for these sentences, but all must involve `I`-final phrase structure.

(150) **Not necessarily `I`-initial**
**Diagnostic: Stranded preposition**

a. *Subject ... Stranded preposition ... V*

þanon he hyre [$_{I'}$ [$_{VP}$ þæne naman **on** _ ] *asette* ]

thence he it the name on set

'From that, he gave the name to it [=the plant *apollinaris*]' (coherbar,Lch_I_[Herb]:23.0.635)

b. *Conjunction ... Subject ... Stranded preposition ... V*

**&** him [$_{I'}$ [$_{VP}$ Godes gast **on** _ ] *wæs* ]

and him God's ghost in was

'And the Holy Ghost was in him' (cootest,Judg:11.1.5719)

Thus, stranded prepositions can be assumed to be reliable diagnostics of both `I`-initial and `I`-final headedness in post- and pre-verbal position respectively.

(iv) **Negatively quantified objects**. Finally, objects that contain a negative quantifier such as *nan* 'no' cannot postpose. They must therefore be analyzed with `I`-initial headedness in post-verbal position. The YCOE annotates the relevant quantifiers with a negative prefix, e.g., `NEG+Q` etc. However, Middle English annotations do not include the `NEG` prefix on quantifiers. I therefore searched for Middle English quantifiers that begin with the letter $n$ as a proxy. Examples of post-verbal negatively quantified objects are given in (151).

(151) **Necessarily `I`-initial**
**Diagnostic: Negatively quantified objects**

a. *Subject ... V ... Negative object*

Se mona [$_{\texttt{I}'}$ *næfð* [$_{\mathsf{VP}}$ **nan** leoht buton of ðære sunnan leoman ]]
the moon not-has no light except of the sun's rays

'The moon does not have any light except through the rays of the sun' (cotempo,ÆTemp:3.7.93)

b. *Conjunction ... Subject ... - V ... Negative object*

**&** heora nan [$_{\texttt{I}'}$ *næfð* [$_{\mathsf{VP}}$ **nænne** leoman buton of ðære sunnan leoman ]]
and of-them none not-has no rays except of the sun's rays

'And non of them has any lights except through the rays of the sun' (cotempo,ÆTemp:1.31.57)

Pre-verbal negatively quantified objects may indicate the position of `VP`. However, like pronouns and marginally particles, they may also prepose. My formal grammar model does not implement this option. Examples of pre-verbal negatively quantified objects are presented in (152). I show with labeled brackets a plausible `I`-final parse for these structures.

(152) **Not necessarily `I`-initial**
**Diagnostic: Negatively quantified objects**

a. *Subject ... Negative object ... V*

swa eac þa gode weorc bute þan rihten geleafen [$_{\texttt{I}'}$ [$_{\mathsf{VP}}$ **nane** mæn ] *ne helpð* ]
so also the good work except the right belief no man not helps

'Likewise, a good deed does not help anybody without the right faith' (coalcuin,Alc_[Warn_35]:19.17)

b. *Conjunction ... Subject ... Negative object ... V*

**and** he [$_{\texttt{I}'}$ [$_{\mathsf{VP}}$ syððan **nan** word ] *ne gecwæð* ]
and he afterwards no word not spoke

'And he didn't say anything afterwards' (ChronC_[Rositzke]:1042.1.1807)

Thus, negatively quantified objects must be analyzed with `I`-initial phrase structure in post-verbal position, and may plausibly, though not necessarily, indicate `I`-final headedness in post-verbal position as well.

The increase in necessarily `I`-initial headedness with each of the diagnostics has been demonstrated before (see section 3.3.1.3). Thus, I saw little reason to conduct the searches separately for different diagnostic elements here, but instead combined them all together. Consequently, I ran a total of 2 (word order) * 2 (period; Old and Middle English) = 4 search queries.[10] The queries for this study were carried out on the original, unrevised corpora. The searches for structures that are not necessarily `I`-initial were executed on the complement file of the results for structures that are necessarily `I`-initial. In this way, clauses involving a diagnostic both before and after the finite verb were counted only once, namely towards the innovative group.

[10]The CorpusSearch query and output files are stored on the Dissertation DVD under `Chapter 3 - Conjunct Clauses/03 Hypothesis Testing/H6`.

**Results**

I fitted a mixed-effects logistic regression to the data measuring the rise of necessarily `I`-initial headedness with diagnostics. It includes time, clause type and their interaction as fixed effects and varying intercepts for every text as a random effect . The result is shown in table 3.12.

```
formula = IInitial ~  Year + ClauseType + Year:ClauseType + (1 | Text),
              family = binomial, data = H6

     Fixed effects:
                                Estimate     Std.Error z-value p
     Intercept                  -7.1874      1.3956    -5.150  <0.001***
     Year                       0.006426     0.001379  4.661   <0.001***
     ClauseType(MC→CC)          1.2314       0.8420    1.463   0.1436
     Year:ClauseType(MC→CC)     -0.001607    0.000840  -1.914  0.0557

     Random effect:
     Text, N=96
     Variance of random intercepts:  1.068

     Null deviance:         8512 on 6575 degrees of freedom
     Residual deviance:     7739 on 6571 degrees of freedom
     AIC: 7749
```

Table 3.12: Logistic Regression Model 4: mixed-effects model for necessarily `I`-initial structures by time and clause type

The model calculates an increase in the log-odds of necessarily `I`-initial structures of 0.0064 per year. The time variable is highly significant, confirming the general increase in `I`-initial headedness during Early English. Neither the clause type variable nor the clause type - year interaction are returned as significant predictors in this model. The lack of statistical significance of the first factor shows that the overall probabilities of finding an `I`-initial structure are quite similar in MCs and CCs and therefore opens up the possibility that the underlying rates in `IP` headedness are not substantially different between these two clause types. The statistical insignificance of the second term suggests that MCs and CCs may innovate the headedness of `IP` at the same rate. Indeed, a likelihood ratio test on the deviance difference between a reduced and a full model also indicates that the inclusion of the time - clause type interaction term does not improve the model enough to justify its inclusion ($\chi^2$=3.35, df=1, $p$=0.06712). Consequently, this finding offers evidence for the prediction made in Hypothesis 6. However, the interaction term only barely misses the customary 5% significance level. The conclusion that MCs and CCs develop at the same rate therefore remains relatively uncertain. The `I`-initial data comes from only 96 texts; there are several texts with neither conservative nor innovative examples of the relative order between finite verbs and diagnostic elements. The texts show a variance in their intercepts of 1.068. This indicates substantial variability among texts. Once again, the inclusion of varying text intercepts proved quite important in the present study. If the random text effect had been omitted, the evidence for a constant rate effect between MCs and CCs would become much stronger. Hence it becomes much more likely that MCs and CCs may develop `I`-initial headedness at different rates when variability among texts is controlled for (though still not sufficient evidence to sustain such a claim). One would thus assign greater certainty to the presence of a constant rate effect than is actually warranted by the data.

I will not evaluate the performance of the model in table 3.12. It would not be a reasonable model to propose for the development of `I`-initial headedness.

The graph in figure 3.6 represents the data I collected for this study. It includes the data points for MCs in green and for CCs in red. The graph also includes regression lines and their 95%-confidence intervals from two independent mixed-effects models for the MC and CC data respectively. They predict the occurrence of an `I`-initial clause from time alone while allowing for random intercepts for each text. In general, the red CC data points fall slightly below the green MC data points. However, the difference does not seem to be very great and the 95%-confidence intervals for the two regression lines are partially overlapping. This shows that the overall probability of finding a necessarily `I`-initial clause is similar in MCs and CCs and the underlying rate may in fact possibly be identical. The two regression lines move quite parallel to each other. This supports the core statement of Hypothesis 6 - the rate of innovating `IP`-headedness is not influenced by the presence of a conjunction and hence develops at the same rate of change in MCs and CCs.

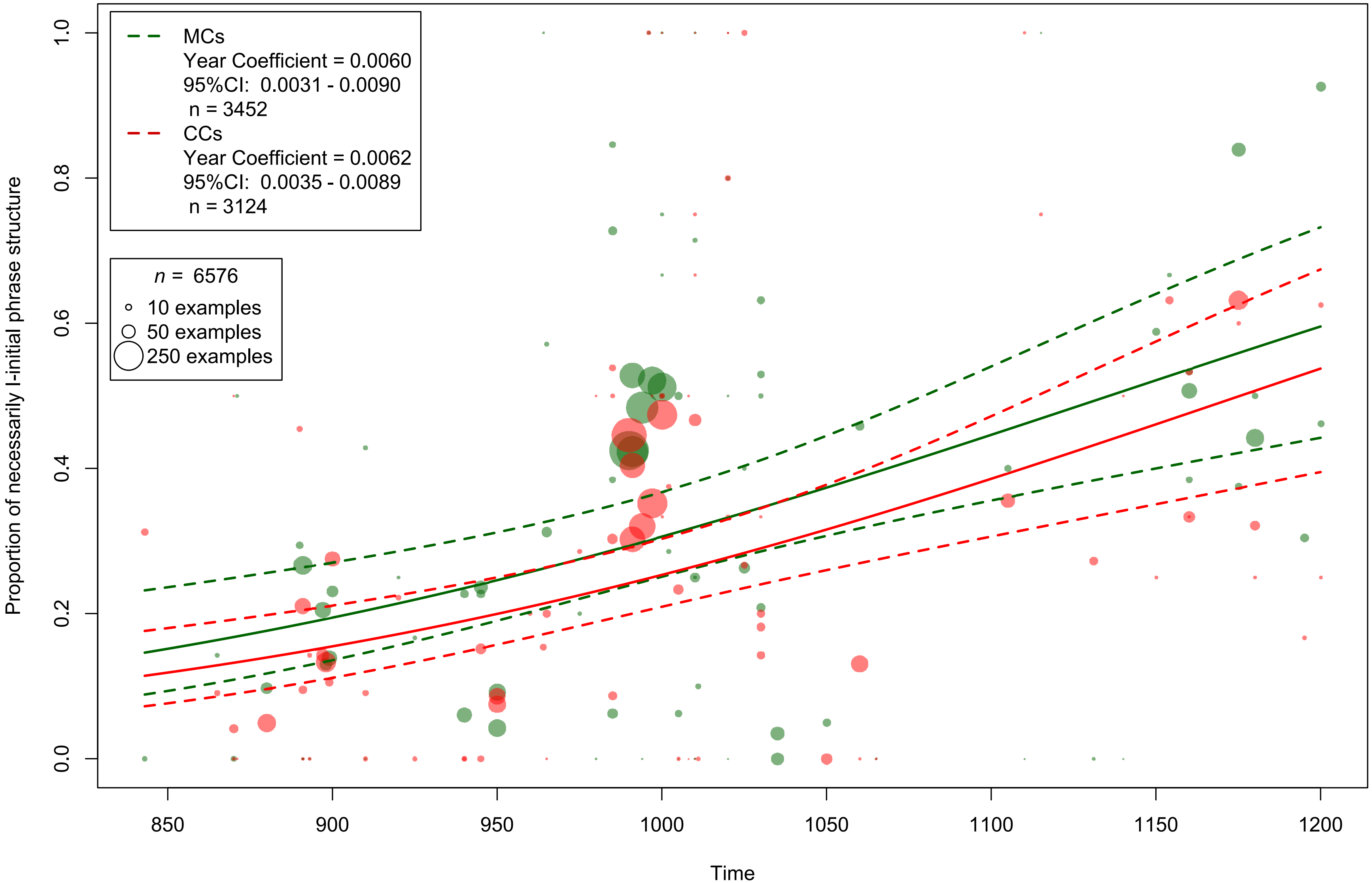



Figure 3.6: Illustration of the development of necessarily `I`-initial headedness

One may find it surprising that my measurement reveals a fairly low proportion of necessarily `I`-initial clauses for the early Middle English era (only c. 50-60%). The reason for this is that non-subject pronouns, the most common of all the diagnostics, still commonly prepose in that period, thus boosting the number of clauses categorized as the counter-context, non-necessarily `I`-initial clauses. A few examples are shown below. The measurement only provides a lower bound of the rate of `I`-initial clauses. The actual rate of `I`-initial phrase structure in early English can be assumed to be substantially higher than what is suggested by figure 3.6.

(153) **Not necessarily `I`-initial clauses with preposed pronouns in early Middle English**

a. Hi **me**reweð swa swiðe ðat ic reste ne mai habben
they me rue so much that I rest not may have
'They cause me so much grief that I cannot have any rest' (CMVICES1,115.1405)

b. & ure Laferrd Jesu Crist / **Himm** ȝaff anndswere
and out Lord Jesus Christ him gave answer
'And our Lord Jesus Christ gave him an answer' (CMORM,II,227.2428)

c. ærost man **hem** beræfoð her eahte
first one them bereaved their property
'First, their property was taken away from them' (CMPETERB1,46.141)

d. and te londes men **hire** bigaleð oðer wile
and the land's men her beguile other while
'And the peasants sometimes charm her' (CMTRINIT,197.2724)

To conclude, the sixth hypothesis can be accepted as well. It is reasonable to assume that MCs and CCs develop `I`-initial headedness at the same rate. Previous research demonstrated the presence of a constant rate effect for the rise of `I`-initial structure between matrix and embedded clauses (see section 3.3.1.3). The present study supplements this finding for the contextual division between MCs and CCs as well. My grammar model predicts this result because the loss of C-head conjunctions does not interact with the rise of innovative `IP`-headedness in any appreciable way.

#### 3.4.3.7 Hypothesis 7: Differential Development of Object Topicalization

Object topicalization becomes less common during the Old and early Middle English periods. My model predicts that this decline should happen faster in MCs than in CCs. The hypothesis can be investigated by regressing the proportion of fronted objects against time and clause type.

**Data Collection**

I measured the rates of topicalization of object `DP`s in the following way. I collected clauses with subject pronouns as indicators of the position of the `IP` boundary and a subsequent finite verb. In addition, the clauses contained an overt nominal object. The object could either immediately precede the pronominal subject, or follow it at an arbitrary distance. I counted MCs separately from CCs. This led to four logically possible word order alignments.

The first structure places the object `DP` immediately before the subject pronoun and does not contain a conjunction. Additional elements may precede the topic. This configuration indicates topicalization in MCs, (154a). The second structure is parallel to the first except that it also includes a conjunction. The conjunction must appear in clause-initial position, but other elements may precede the topic category. This word order pattern indicates topicalization in CCs, (154b). The third construction type includes an object positioned after the finite verb. That is to say, there is an object lower in the structure that could theoretically have fronted but failed to do so (abstracting away from information structure constraints). The clause must be introduced by the subject pronoun in first position. This word order type reflects MCs in which the topicalization rule did not apply, (154c). The last configuration also shows an object that could potentially have topicalized to the high topic position but instead remained lower in the structure. Now, however, the clause-initial position must show a conjunction followed immediately by the pronominal subject. Such structures reflect CCs with no topicalization, (154d).

(154) a. **Topicalization**
**+MC**
*Object - Pronominal subject ... V*

[DP Langsume tale ] we magon macian be ðysum
long tale we may make about this

'We could write a long tale about this'
(coaelhom,ÆHom_23:80.3745)

b. **Topicalization**
**+CC**
*Conjunction ... Object - Pronominal subject ... V*

**&** [DP horses hyda ] hi habbað him to hrægle gedon
and horse's hide they have themselves to clothing done

'And they used horse hide for their clothing'
(comarvel,Marv:26.1.130)

c. **No topicalization**
**+MC**
*Pronominal subject ... { V, Object }*

He arærde ða on ðære ylcan byrig [DP mære cyrcan ]
He reared then in the same city great church

'He then built a great church in the same city'
(cocathom2,ÆCHom_II,_38:287.262.6496)

d. **No topicalization**
**+CC**
*Conjunction - Pronominal subject ... { V, Object }*

**Ac** he worhte [DP fela wundra ] ætforan þam deman
but he worked many wonders before the judge

'But he performed many miracles in front of the judge'
(coaelive,ÆLS_[Exalt_of_Cross]:202.5672)

Every example was coded for its source text file and its approximate date of composition. In total I conducted 2 (word order) * 2 (clause type) * 2 (period; Old and Middle English) = 8 search queries.[11]

**Results**

I calculated the rate of topicalization as the proportion of sentences with a topic in pre-subject position out of all relevant sentences. Table 3.13 summarizes the results of this study.

| Clause type | Topicalization *Object - spro ... V* | No Topicalization *spro ... { V, Object }* |
|---|---|---|
| MC | 1063 (31.6%) | 2,306 (68.4%) |
| CC | 579 (16.4%) | 2,944 (83.6%) |

Table 3.13: Object topicalization rates in MCs and CCs

As expected, the rate of object topicalization is significantly lower in CCs, 16.4%, than in MCs, 31.6%, and the size of the clause type effect, as measured by the odds ratios, is moderately large ($\chi^2$=216.0, df=1, $p$<0.001, odds ratio = 2.34, 95%-confidence intervals [2.09 – 2.63). This confirms the claim that the difference between MCs and CCs has distributional consequences not just for verb placement but also for topicalization.

I then fitted a mixed-effects logistic regression model to the data. It estimates, once again, the realization of the dependent word order variable from time, clause type and their interaction as fixed effects and includes varying intercepts for every text as a random effect. The resulting model is presented in table 3.14.

[11] The CorpusSearch query and output files can be found on the Dissertation DVD under `Chapter 3 - Conjunct Clauses/03 Hypothesis Testing/H7`.

```
formula = ObjTopic ~  Year + ClauseType + Year:ClauseType + (1 | Text),
              family = binomial, data = H7

         Fixed effects:
                                   Estimate     Std.Error z-value p
         Intercept                 1.7563       1.0286    1.707   0.0877
         Year                      -0.0024608   0.0010169 -2.420  0.0155*
         ClauseType(MC→CC)         -5.2574      0.7918    -6.640  <0.001***
         Year:ClauseType(MC→CC)    0.0044217    0.0008004 5.524   <0.001***

         Random effect:
         Text, N=96
         Variance of random intercepts:  0.5307

         Null deviance:         7568 on 6891 degrees of freedom
         Residual deviance:     7053 on 6887 degrees of freedom
         AIC: 7063
```

Table 3.14: Logistic Regression Model 5: mixed-effects model for object topicalization by time and clause type

The model estimates that object topicalization in MCs declines at a rate of -0.0025 log-odds per year. At the same time, the incidence of object topicalization in CCs is predicted to increase by -0.00246 + 0.00442 = 0.00196 log-odds per year. This suggests that the loss of C-head conjunctions not only leads to a differential reduction in topicalization between MCs and CCs, but compensates for the declining rate of object topics to such a degree that the former change actually reverses the surface effects of the latter. This over-compensation may make the global decline in topicalization hard to recognize for the model so that the time variable is only significant at the 5% level. The interaction term measuring the divergence in slopes between MCs and CCs, however, is a highly significant predictor in this model. A likelihood ratio test on the deviance difference between a reduced and a full model also necessitates the inclusion of the time - clause type interaction term ($\chi^2$=23.21, df=1, $p$<0.001). These findings support the hypothesis that the reduction in object topicalization does not proceed at a constant rate between MCs and CCs but occurs more rapidly in the former than the latter clause type. As the clause type category is changed from MC to CC, the probability of finding a topicalized object is lowered significantly by log-odds of -5.26. This represents the fact that object topicalization is far more likely to occur in MCs than in CCs. The data is taken from 96 different text files whose random intercepts vary by about 0.5307, which indicates only moderate variability among texts.

The model has a poor quality. It fits significantly better than a null model (where the null model is an intercept only model) (Likelihood Ratio Test, $\chi^2$= 703.1, df=4, $p$<0.001), but still fits the data badly overall (Pseudo-$R^2_{marginal}$= 0.053, Pseudo-$R^2_{conditional}$=0.185). The model has almost no classificatory power above the baseline (C-index=0.712, classification accuracy =76.8% vs. baseline: 76.2%). There may be three reasons for the poorness of this model. (i) As before, models representing linguistic changes from the middle ages tend to perform poorly because of low data quality generally. (ii) The cross-over interaction between time and clause type may not adequately model the actual time course of the development. For instance, it is conceivable that topicalization declines during Old English, albeit at a slower rate in CCs than MCs, and increases in CCs only towards the early Middle English period. One would then have to model a u-shaped curve for the development in CCs. (iii) The reduction in topicalization may be poorly implemented in terms of competition between an innovative and a conservative form. Rather than thinking of the change in terms of the loss of a word order pattern, one should conceptualize it as a change in the information-structural conditions on licensing this word order alignment. The model may somehow reflect the improper measurement that leaves information structure unconsidered. Since I cannot implement practicable solutions for any of these problems without going beyond the technical scope of this chapter, I shall accept the logistic regression in table 3.14 as a possible model for the reduction in object topicalization in Early English despite its obvious shortcomings.

The model is illustrated in figure 3.7. The graph depicts the proportion of topicalized objects in each of the text files depicted with green points for MCs and with red points for CCs. In general, the red points fall below the green ones, which highlights the lower probability of object topicalization in CCs than in MCs. The plot also includes the regression lines for MCs and CCs and their 95%-confidence intervals. The green MC regression line slopes downward whereas the red CC curve moves up. This crossing pattern illustrates the hypothesized differential development of object topicalization. By about 1200, the probabilities of finding a topicalized object in MCs and CCs have becomes so similar that there remains essentially no evidence for the originally dissimilar distribution of topicalized objects between these two clause type.

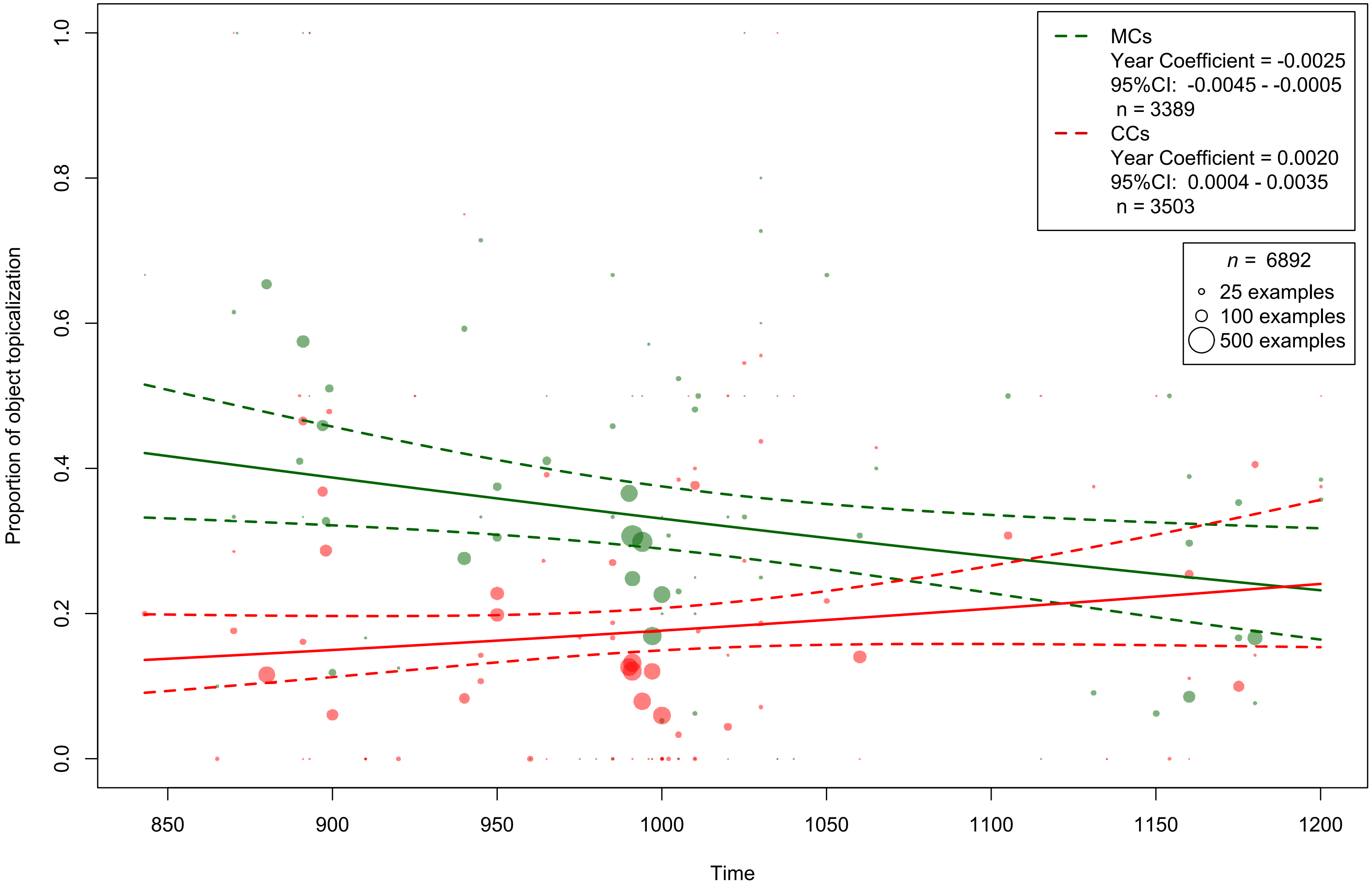



Figure 3.7: Illustration of the mixed-effects model for the development of object topicalization

My findings are very much in accordance with speculations about the time course of topicalization expressed by Kohonen on the basis of just two Middle English texts. He mentions the following in passing:

> The *and/ac* clauses had an interesting increase of topicalization in V[ices and]V[irtues] and S[awles] W[arde] [two early Middle English texts]. [...] [T]his could be interpreted as a change in the status of the *and* and *ac* conjunctions: in OE, these conjunctions had the effect of blocking topicalization and causing a dependent clause word order (SXV), i.e., they shared properties of subordinating conjunctions. It seems possible that the change took place by about 1200 [...]; after this, clauses introduced by these conjunctions had increasingly similar features to other main clauses. (Kohonen 1978: 154)

The results of my measurement of direct object topicalization agree with Kohonen's contention that (i) CCs may increase the rate of topicalization during the early Middle English period and that (ii) by about 1200, the rates of topicalization are no longer substantially different between MCs and CCs. Kohonen also outlines an explanation for these findings in terms that are similar to my concept of C-head conjunctions. He says that conjunctions "block" topicalization, which may be synonymous to disabling the `C`-layer for topicalization, and even sees a direct relation between a reduction in topicalization and distributional differences in verb placement between CCs and MCs as indicated by the use of the word "cause." I would suggest that Kohonen's formulation of this informal marginal note should be regarded as both insightful for its time as well as essentially correct.

However, my Early English grammar model also revises some central aspects of Kohonen's account. First, Kohonen's measurement of topicalization lacks precision. He does not distinguish between different kinds of topic categories, but counts all clause-initial non-subjects as topics, including nominal objects, `PP` and `ADVP` adjuncts, pronominal objects and even sentential negation. He also jointly considers all kinds of subjects. My measurement is more restrictive and, I would argue, more accurate. Second, I do not agree that Old English conjunction should be said to share "properties of subordinating conjunctions." Rather some Old English conjunctions and complementizers share a syntactic position. Third, *SXV* patterns should not be described as "dependent word order" because `I`-final and `I`-initial phrase structure are available in both matrix and embedded clauses.

In conclusion, while the regression model proposed here is of a low quality, the trend in the development of object topicalization is quite clear. MCs do indeed seem to reduce the incidence of topicalized objects at a faster rate than CCs. In fact, there appears to be a slight increase in object topicalization in the latter clause type. One is thus justified in accepting Hypothesis 7. This in turn lends further plausibility to my grammar model, as it successfully explains the observed diachronic divergences in topicalization between MCs and CCs as the interaction between the gradual loss of conjunctions that can be placed under `C` and the simultaneous reduction in object topicalization.

#### 3.4.3.8 Hypothesis 8: Differential Development of Subject Topicalization

My final hypothesis states that potential subject topics should be less common in CCs than MCs. Furthermore, the frequencies of potentially topicalized subjects should converge between those two clause types over time. I will test this hypothesis by collecting clauses with non-subject pronouns as possible markers of the `CP`-`IP`-boundary and record the occurrence of full subjects either in front of or after them. One can then track the development of 'full subject - pronoun' structures as a proxy for potential subject topicalization.

**Data Collection**

It is not an easy task to ascertain whether a full subject is placed in the clause-initial topic position or not. The reason is that my grammar model allows for the analysis of full subjects in 'subject - verb' sentences as an element both in the default subject position, the specifier of `IP`, as well as in the more prominent topic position, the specifier of `CP` (see (74)). As a side note, the structural ambiguity is reminiscent of the positional indeterminacy of the *wh*-pronoun in Modern English subject questions, such as *Who loves Mary?*. I therefore decided to collect sentences that may potentially involve the application of the subject topicalization rule and to compare them to structures that unequivocally do not involve it.

Clauses that have a relatively high probability of containing topicalized subjects include a clause-initial full subject immediately followed by a pronoun. If the pronoun is parsed with the rule for pronominal preposing to `IP`, (81), a preceding nominal subject must have topicalized to the specifier of `CP`, [$_{\texttt{CP}}$ `subject` [$_{\texttt{IP}}$ `pro ... V` ]]. However, the preposed pronoun may also occur at `I′` by rule (79). In this case, subject topicalization may, [$_{\texttt{CP}}$ `subject` [$_{\texttt{IP}}$ [$_{\texttt{I′}}$ `pro ... V` ]]], or may not be involved, [$_{\texttt{IP}}$ `subject` [$_{\texttt{I′}}$ `pro ... V` ]]. If the pronoun does not prepose but occurs in the ordinary positon for objects within the `VP` by rule (23a), forcing `I`-final structure and, where necessary, constituent extra-position, the clause likewise may or may not involve subject topicalization, [$_{\texttt{CP/IP/VP}}$ `subject` [$_{\texttt{V′}}$ `pro` ] `V` ]. The last parse is not very likely because `I`-final phrase structure is not very common in main clauses. There are also attestations of post-verbal diagnostics in such structures, which unequivocally require `I`-initial structure and hence preposing of the pronoun out of the `VP`. What is important is that all of these analyses for the ‘full subject - pronoun’ pattern open up the potential for topicalization of the subject.

I searched for such structures as follows. I required that a full subject should immediately precede a pronoun. I did not restrict the position of the full subject within the clause any further, i.e., the subject could be preceded by other constituents, say `CP`-adjoined sentence adverbs, or adverbial subordinate clauses, etc. The finite verb followed the subject at any distance. Examples are shown in (155) for MCs and in (156) for CCs. The a.-examples are typical of my data set. The b.-examples additionally include a post-verbal diagnostic, thus forcing `I`-initial structure and pronominal preposing to the `I`-domain. I indicate with labeled bracketing the targeted analysis with pronominal preposing to `IP` and subject topicalization to the specifier of `CP`.

(155) **Potentially topicalized full subject**
**+MC**

a. *Full subject - Pronoun ... Verb*

[$_{\texttt{CP}}$ God [$_{\texttt{IP}}$ **him** *wearð* þa yrre ]]
God him became then angry

‘God then became angry with him’ (coaelive,ÆLS[Pr_Moses]:225.2983)

b. *Full subject - Pronoun ... Verb ... Diagnostic*

[$_{\texttt{CP}}$ Se hælend [$_{\texttt{IP}}$ **him** [$_{\texttt{I′}}$ *cwæð* to _, Ic beo sylf mid þe ]]]

the savior him spoke to I am self with you

‘The Savior said to him, “I myself am with you.”’ (coaelive,ÆLS_[Julian_and_Basilissa]:16.942)

(156) **Potentially topicalized full subject**
**+CC**

a. *Conjunction ... Full subject - Pronoun ... Verb*

**and** [$_{\texttt{CP}}$ Apollonius [$_{\texttt{IP}}$ **hi** *bæd* ealle gretan ]]
and Apollonius them bade all greet

‘And Apollonius ordered them all to salute’ (coapollo,ApT:11.4.184)

b. *Conjunction ... Full subject - Pronoun ... Verb ... Diagnostic*

**ond** [$_{\texttt{CP}}$ misenlico wilddeor [$_{\texttt{IP}}$ **him** [$_{\texttt{I′}}$ þær *comon* to _ ]]]

and diverse wild-animals him there came to
‘And various animals came to him there’ (comart3,Mart_5_[Kotzor]:Ju2,A.6.887)

In contrast, clauses must necessarily be parsed without subject topicalization if they show a pronoun followed by a full subject at any distance. The pronoun may have topicalized to the specifier of `CP` by rule (67). I do not think that such a parse is particularly likely due to the low deictic value of personal pronouns, but since most scholars believe that object pronoun topics are a genuine grammatical option in Old English, my model allows for the generation of such parses, [$_{\texttt{CP}}$ `pro ...` [$_{\texttt{IP}}$ { `subject, V` } ]]. Alternatively, the pronoun may have preposed to `IP` by rule (81), [$_{\texttt{IP}}$ `pro` [$_{\texttt{IP}}$ `...` { `subject, V` } ]], or to `I′` by rule (79), [$_{\texttt{I′}}$ `pro ...` { `subject, V` } ]. Crucially, none of these parses of ‘pronoun - full subject’ patterns permit the full subject to have topicalized.

Structures with impossible subject topicalization were retrieved as follows. I looked for clauses in which the pronoun was the first constituent in the clause for MCs, and in which the conjunction was positioned first with the pronoun immediately following for CCs. Both the full subject and the finite verb had to follow the initial pronoun, but otherwise the orders between pronoun and full subject or between pronoun and verb were left unrestricted. The examples in (157) illustrate for MCs, the examples in (158) for CCs. The a.-examples show the order in which the full subject precedes the verb. The b.-sentences exemplify the alternative word order with the verb preceding the full subject. The c.-examples additionally involve a post-verbal diagnostic constraining the `IP` to be head-initial. Because of the substantial structural ambiguity of the structures under consideration here, I do not indicate any particular position for the fronted pronouns with labeled brackets.

(157) **Impossible full subject topicalization**
**+MC**

a. *Pronoun ... { Full subject, Verb }*

**Hine** [ se ercebiscop mid his agene hond on horse *ahof* ]
him the archbishop with his own hand on horse up-heaved

'The archbishop lifted him up with his own hand onto the horse' (cochad,LS_3_[Chad]:49.31)

b. *Pronoun ... { Full subject, Verb }*

**hit** *hæfde* [ Agustus him to onwalde geseald ]
it had Augustus him to power given

'Augsuts had given it [=Egypt] to him to rule over' (coorosiu,Or_6:1.134.2.2821)

c. *Pronoun ... { Full subject, Verb } ... Diagnostic*

**Hire** [$_{I'}$ *com* [$_{VP}$ ælce niht sumes haliges wifes gast to _ ]]
her came each night of-some holy woman ghost to
'The spirit of a holy woman came to her each night'
(comart1,Mart_1_[Herzfeld-Kotzor]:De25,B.10.26)

(158) **Impossible full subject topicalization**
**+CC**

a. *Conjunction - Pronoun ... { Full subject, Verb }*

**And eow** [ unwæstm þurh unweder gelome *gelimpeð* ]
and to-you un-fruit through un-weather often happens

'And bad harvest will frequently happen to you because of bad weather'
(cowulf,WHom_19:64.1530)

b. *Conjunction - Pronoun ... { Full subject, Verb }*

**& hine** *ne mihte* [ nan mann mid racenteagum gehealdan ]
and him not might no man with chain-cords hold

'And no man could hold him with fetters' (coaelhom,ÆHom_18:221.2606)

c. *Conjunction - Pronoun ... { Full subject, Verb } ... Diagnostic*

**& him** [$_{IP}$ Scipia [$_{I'}$ *sende* [$_{VP}$ sciphere æfter _ ]]]
and them Scipio sent ship-army after
'And Scipio sent a fleet after them' (coorosiu,Or_4:10.106.31.2216)

I recorded the source text file and its approximate date of composition for every example. In total, I carried out 2 (word order) * 2 (clause type) * 2 (period; Old and Middle English) = 8 search queries.[12]

[12]The CorpusSearch query and output files can be found on the Dissertation DVD under `Chapter 3 - Conjunct Clauses/03 Hypothesis Testing/H8`.

**Results**

I determined the ratio between clauses in which subject topicalization might have applied to those in which subject topicalization was not a possibility. The result of this calculation is presented below in Table 3.15.

| Clause type | Subject topicalization possible *Full subject - Pronoun ... Verb* | Subject topicalization impossible *Pronoun ... {Verb, Full subject}* |
|---|---|---|
| MC | 893 (81.2%) | 207 (18.8%) |
| CC | 787 (62.6%) | 470 (37.4%) |

Table 3.15: Potential subject topicalization in MCs and CCs

As expected, the potential for subject topicalization is significantly lower in CCs than MCs. The association between clause type and potential topicalization in this sample is comparable to the measurements of object topicalization as shown by a similar value of the odds ratio ($\chi^2$=97.9, df=1, $p$<0.001, odds ratio = 2.57, 95%-confidence intervals [2.12 – 3.13]). Hence, this investigation, too, provides some evidence for a lower rate of topicalization in CCs than in MCs.

To repeat, my model accounts for this distributional difference as follows. Full subjects in MCs are always allowed to topicalize to the specifier of CP across a preposed pronoun. They will therefore often appear in front of pronouns.

(159) [$_{CP}$ Se hælend [$_{IP}$ **him** [$_{I'}$ *cwæð* to _ ]]]
the savior him spoke to
'The Savior said to him ... ' (coaelive,ÆLS_[Julian_and_Basilissa]:16.942)

In contrast, subjects will not be able to topicalize if a C-head conjunction blocks the C-domain. They will therefore often be placed in post-pronominal position in CCs.

(160) [$_{CP}$ **&** [$_{IP}$ **him** Scipia [$_{I'}$ *sende* sciphere æfter _ ]]]
and them Scipio sent ship-army after
'And Scipio sent a fleet after them' (coorosiu,Or_4:10.106.31.2216)

This mechanism results in the observed higher proportion of clauses without potential subject topicalization in CCs than MCs. However, pre-pronominal subjects are not entirely banned from the former clause type because alternative parses for 'conjunction - full subject - pronoun' patterns remain available, for instance, involving full subjects in the specifier of IP with preposed pronouns to I′ and C-head conjunctions, or topicalized subjects in the specifier of CP and high logical connectors.

Incidentally, 'pronoun - full subject' clauses have been used to argue for the existence of verb-third structures with a topic in the specifier of CP and a full subject in the specifier of IP (e.g., Speyer 2010: 199). The sentence in (161) is a notable example. It involves a clause initial pronoun as the presumed topic category, a full subject and the post verbal particle *ongean* 'again.'

(161) **and him** [$_{IP}$ se innoþ eac [$_{I'}$ *geopenode* [$_{VP}$ ongean ]]]
and him the bowels also opened again
'And his bowels also opened again' (coaelive,ÆLS_[Vincent]:170.7907)
(cited as a V3 sentence in Speyer 2010: 199, ex. (17b)).

The intended parse is in fact a possible analysis in my model. However, in light of the word order distribution reported in table 3.15, it seems important to note that this clause features an initial conjunction. Hence, it seems quite possible for the structure to be analyzed with a pronoun that has preposed to the IP-CP-boundary and a full subject that failed to topicalize because the C-layer has been rendered inactive by a C-head conjunction. The clause does not necessarily evidence a verb-third sentence of the intended kind under my model assumptions.

More importantly, some scholars have presented evidence for different rates of subject-verb inversion in MCs and CCs with topicalized objects. For instance, Dreschler (2015: 237) finds that 64.6% of MCs show inversion of verb and full subject after object topics, 'object - verb - full subject', but only 54.7% of CCs show subject verb inversion in this context, 'conjunction ... object - verb - full subject.' If true, this observation would be unexpected under my model and hence challenge its validity. However, Dreschler's counts include a large number of supposed topicalized object pronouns, i.e., structures such as (161). Under my model, the pronoun may not be a topic, but be placed inside the `I`-domain. The initial conjunction might prevent fronting of the subject across the pronoun to the specifier of `CP`. Therefore, many CCs will surface with the order ‘pronoun - full subject,’ elevating the number of supposedly uninverted sentences, where the actual contrasting MCs would show the order ‘full subject - pronoun,’ but were ignored in the investigation. Dreschler’s observation does therefore not directly undermine confidence in the correctness of my grammar model.

Next, I fitted a mixed-effects model to the potential subject topicalization data. It predicts, once more, the word order variation from time, clause type, their interaction and a random text effect. The resultant model is given in table 3.16.

```
formula = SubjTopic ~  Year + ClauseType + Year:ClauseType + (1 | Text),
             family = binomial, data = H8

     Fixed effects:
                               Estimate   Std.Error   z-value   p
     Intercept                 -3.8622    1.3472      -2.867    0.0041**
     Year                      0.004506   0.001350    3.337     <0.001***
     ClauseType(CC→MC)         5.3591     1.2179      4.400     <0.001***
     Year:ClauseType(CC→MC)    -0.004561  0.001232    -3.702    <0.001***

     Random effect:
     Text, N=84
     Variance of random intercepts:  0.4765

     Null deviance:        2827 on 2356 degrees of freedom
     Residual deviance:    2607 on 2352 degrees of freedom
     AIC: 2617
```

Table 3.16: Logistic Regression Model 6: mixed-effects model for subject topicalization by time and clause type

I report the model with CCs as the reference value for the clause type category rather than with MC as its reference as before. This notation enables a somewhat more convenient discussion. The model asserts that CCs develop structures that may involve subject topicalization at a rate of 0.0045 log-odds per year. The increase is highly significant. Overall, potentially topicalized subjects are by about 5.4 log-odds more likely in MCs than CCs. The difference is highly significant in this model as well. In contrast, the proportion of potentially topicalized subjects in MCs remains essentially unchanged throughout the Early English period, as shown by a rate of change close to zero, 0.00451 - 0.00456 = -0.00005 log-odds per year. This difference in the rates of change between MCs and CCs, too, is a highly significant predictor in this model. The year - clause type interaction term should be included in the model as shown by a likelihood ratio test on the deviance difference between a reduced and a full model ($\chi^2$=6.52, df=1, $p$=0.0107). These findings instantiate one possible diachronic pattern that supports the present hypothesis: The majority of possible subject topicalization occurs in MCs, which remain stable, whereas potentially topicalized subjects are less common in CCs, but are on the rise in this clause type. Consequently, the logistic regression does indeed conform to the expectation of finding a convergence between the two clause types. The data for Hypothesis 8 comes from 84 text files with relatively little random variability as shown by a low value of the variance parameter of just 0.4765.

Once again, the model performance is quite weak. While it fits significantly better than a null model (where the null model is an intercept only model) (Likelihood Ratio Test, $\chi^2$= 322.5, df=4, $p$<0.001), the overall data fit is poor (Pseudo-$R^2_{marginal}$= 0.059, Pseudo-$R^2_{conditional}$=0.178). The model does not classify substantially more data points correctly than the baseline (C-index=0.722, classification accuracy =74.2% vs. baseline: 71.2%). The deficiencies of the model can be attributed to the same reasons as outlined in the previous studies.

The graph in figure 3.8 represents the mixed-effects model. It depicts the proportion of ‘full subject - pronoun’ orders in every text file in green for MCs and in red for CCs. The red points tend to fall below the green ones, thus illustrating the lower overall probability of potentially topicalized subjects in CCs. The graph also includes regression lines and their 95%-confidence intervals for the developments in MCs and CCs. While the green line for MCs is essentially flat, the red curve for CCs clearly slopes upwards. This wedge-shaped pattern indicates the gradual convergence in potential subject topicalization between the two clause types. Indeed, by about 1200 all evidence for the originally differential distribution of ‘full subject - pronoun’ structures has disappeared. The graph also shows that the data points become smaller and thin out from early Middle English on, which attests to the the onset of the loss of pronominal preposing.

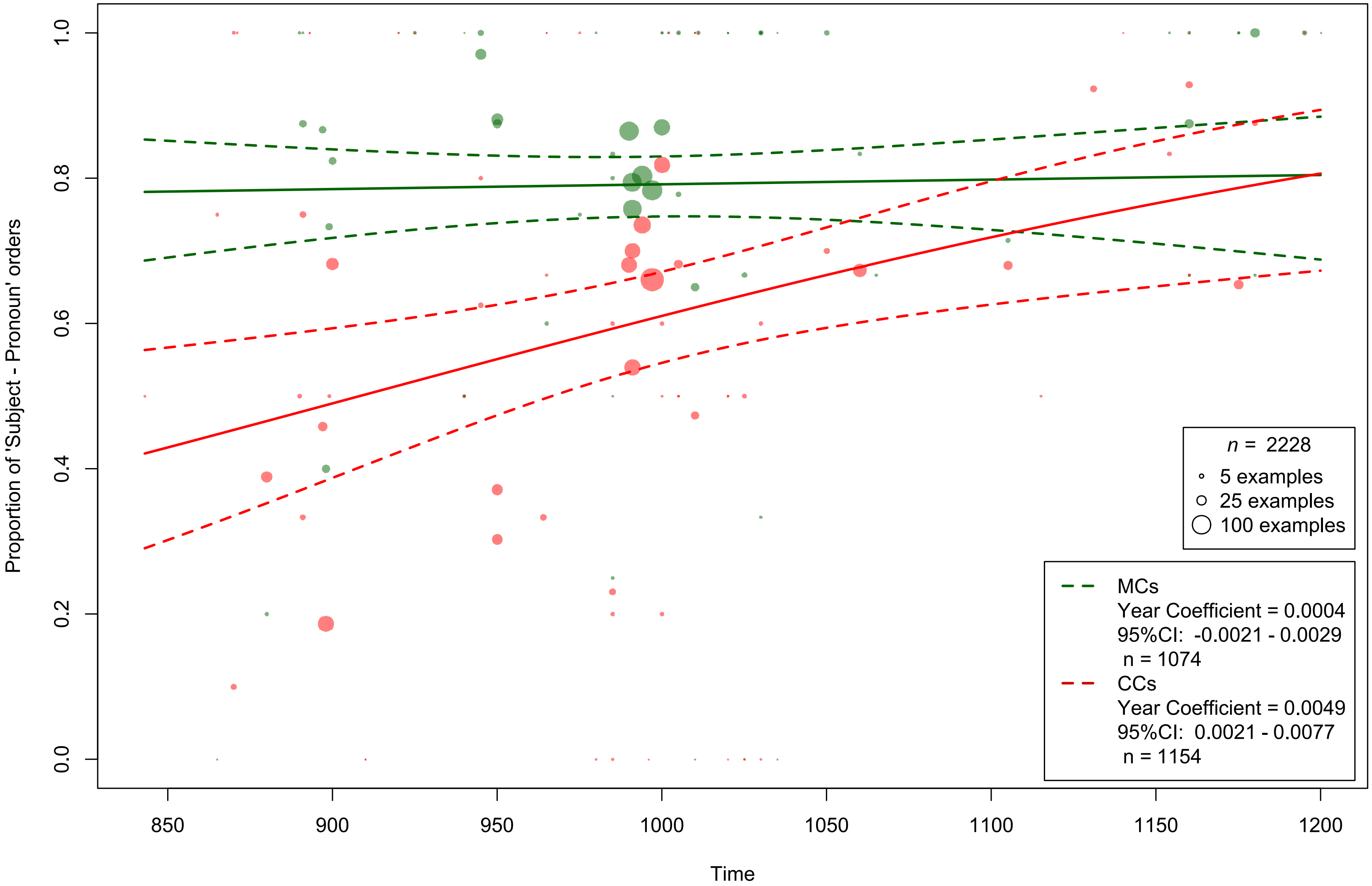



Figure 3.8: Illustration of the mixed-effects model for the development of potential subject topicalization

In summary, I could provide evidence in support of the eighth hypothesis so that one is justified in accepting it. Although it is difficult to measure the rate of subject topicalization directly, a comparison between structures that at least plausibly involve topicalized subjects with word orders that cannot possibly involve them revealed that CCs can be assumed to instantiate substantially fewer structures with subject topicalization than MCs. Further, the proportions of potential subject topics are harmonized over time. Once again, these positive results should strengthen one's confidence in the grammar model proposed here, as it is able to explain the findings as a consequence of a blocking effect induced by C-head conjunctions and the gradual disappearance of this syntactic category.

### 3.4.4 Summary

This part was concerned with tests of eight predictions regarding distributional and diachronic differences between MCs and CCs that were deduced earlier from my theoretical grammar model. I could provide empirical evidence for all of them. Consequently, my grammar model formalizing Old English conjunct clauses, the linguistic changes it incorporates and the proposed interactions between these changes all have received substantial support during the course of these hypothesis tests.

Specifically, I propose that conjunctions can sometimes be inserted under `C` but gradually lose the syntactic category that enables this placement. It follows from this that CCs display fewer structures with the verb under `C` and more clauses with the verb inside the `IP`, and as a side-effect thereof, more manifestations of `I`-final structures (Hypothesis 1), that CCs and MCs may actually have identical, underlying rates of `I`-final and `I`-initial headedness (Hypothesis 2), that the loss of C-head conjunctions will compensate for and seemingly slow down the loss of `V`-to-`C` patterns (Hypothesis 3), that not all conjunctions, but only those immediately placed before `IP`s, are responsible for the special characteristics of CCs (Hypothesis 4), and that the loss of C-head conjunctions will introduce an additional source for the loss of `I`-final structures, and hence speed up the loss of verb-final patterns on the surface, by comparison with all other clause types (Hypothesis 5). One verb placement development that should not be affected by the presence or absence of C-head conjunctions is the rise of `I`-initial structures measured with post-verbal diagnostics. Indeed, CCs and MCs do not innovate `I`-initial phrase structure at significantly different rates of change (Hypothesis 6). Finally, my model assumes that the loss of C-head conjunctions interacts with the decline in topicalization and thus accounts for the observed faster decrease of object topics in MCs than CCs (Hypothesis 7) and for a harmonization in the manifestation of potential subject topics between the two clause types (Hypothesis 8).

I found that by about 1200, texts no longer record great amounts of differential distributions between CCs and MCs in verb placement and topicalization. As a consequence, CCs disappear as a particular, detectable clause type from the language as well.

## 3.5 Discussion

In this chapter, I addressed the research question of whether syntactic changes can interact with one another in such a way as to cause diachronic divergences that are observable in corpus data. The empirical material for this investigation came from the loss of the special status of Old English main clauses introduced by a conjunction, and two additional changes that occur independently at the same time - the rigidification of finite verb placement in declaratives to exclusively verb-medial position, and a decline in argument topicalization. I showed that a computationally explicit, syntactic model in which Old English conjunctions may vie for a structural slot with high, finite verbs can successfully capture the distributional differences regarding verb placement and constituent fronting between MCs and CCs and that this model can also predict measurable interactions between the three linguistic changes in the form of diachronically divergent patterns. Indeed, I subsequently observed violations of constant rate effects between MC and CC contexts in the decline of high verb placement, in the loss of verb-final patterns, and in the reduction of argument topicalization. These findings count as the first, or are among the first, theoretically predicted demonstrations of constant rate effect violations, thereby further substantiating the correctness of the Constant Rate Hypothesis. They thus allow an affirmative answer to my general research question.

As shown in the introductory quote of this chapter, physicist Heinz Pagels suggested in the 1980s that the future would see the investigation of complex systems that are made up of many components interacting with each other. Examples comprise climatology, neurobiology or organizational economics. Historical linguists have provided good descriptions of many different instances of language change in isolation. Perhaps the field can now begin to follow the trend of studying complex systems too. I am confident that the methodology used here to examine interactional syntactic changes could be applied to many other cases and to complex changes recorded in much larger bodies of data.

I would like to think that this study could also contribute in a fruitful way to the study of Old English syntax as such. It leads to a clearer understanding of the distributional differences between MCs and CCs outlined at the beginning of the chapter. However, the theory of C-head conjunctions presented here is not uncontentious. Instead, it competes with alternative explanations for the special status of Old English CCs.

Firstly, there are a number of functional, particularly information-structural explanations for its existence. I have argued that, while these approaches offer many insightful observations and point towards fruitful research questions, they do not offer primary explanations. I would like to reiterate my dissatisfaction regarding functional explanations with an analogy from biology: One could liken a question such as 'why do eagles have wings' to the syntactic problem at hand, 'why does Old English have special conjunct clauses.' There are two acceptable answers. Either one attempts to build a model of eagle feathers drawing on their genotype and its phenotypic variations or one attempts to trace its phylogenetic history. The first explanation is parallel to the construction of a linguistic competence model taking into account why performative realizations in text documents may not correspond perfectly to the posited rules. This is the approach I took in this chapter. The second explanation would be comparable to identifying linguistic principles of fitness, rule selection and so on. This kind of *why*-question is not fully within the scientific reach of historical linguistics at the moment. What one should not do, in my opinion, is to derive from the function of a structure a reason for the existence of the structure. One cannot say that eagles have wings because they want to fly. Likewise, one cannot say that Old English has special conjunct clauses because speakers want to have an option to strongly elaborate, modify or connect a proposition to a preceding one (Bech 2001) etc. Such reasoning employs backwards causation and is incompatible with algorithmic modeling. Teleological explanations of this sort are therefore hard to test empirically, and, indeed, functional accounts of Old English conjunct clauses have, as far as I know, never deduced a quantitatively testable prediction. There is, however, value in studying the manifestation of biological or linguistic systems in all its manifestations. Hence, it is as interesting a research question to investigate how eagles actually do fly as it is to ask how speakers actually do use conjunct clauses. I want to make it clear that I see a lot of value in research on the information-structural properties of Old English constructions. I just do not think that such research can elucidate the cause for the structures it scrutinizes.

A far more serious alternative to the model I offered here proposes that the special status of Old English CCs arises from the coordination of `IP`s (Haeberli 2001). Since `IP`-coordination will cut off the `CP`-layer from the second conjunct, the distributional patterns described in this chapter would follow. In a sense then, the two models are very similar to each other. Arguably, the `IP`-coordination account is more elegant and has the advantage of avoiding the postulation of overt `C`-elements in main clauses. On the other hand, it loses elegance for the description of the diachronic observations. One would have to claim that `IP`-coordination was somehow lost even though it clearly still exists in Modern English (*Why did [$_{IP}$ A do B] and [$_{IP}$ C do D]?*, *I know that [$_{IP}$ A did this] and [$_{IP}$ B did that].* etc.). Further, the change in syntactic category or feature endowment of conjunctions seems to instantiate a frequent kind of development and can easily be formalized (see e.g., the change in possessive *have* in the first chapter). In contrast, the required reduction in `IP`-coordination would remain opaque and badly formalized. The concept of C-head conjunctions can also account more easily for potentially wedged conjunctions that may occur between initial `CP`-adjuncts and the rest of the clause than the alternative proposal of `IP`-coordination.

Thus, the formalization of Old English CCs presented in this chapter not only receives support from empirical evidence adduced during the hypothesis testing process, but also successfully stands up to competing accounts and perspectives. It should therefore be considered as a serious contender for the true explanation of the special status of Old English conjunct clauses.

Despite the relative success of my model, there remains great uncertainty about its adequacy. Skepticism towards the findings of the research presented here is warranted because of imprecise text dating methods, the low data quality of texts from medieval languages, poor performance of the statistical models, and, perhaps most importantly, the fact that the proposed C-head conjunctions are only indirectly observable by their effects on certain word order distributions. In light of these difficulties, there is a chance that my statistical models may inadequately reflect the reality of certain Old English grammatical developments. Furthermore, there are numerous potential confounding factors that I did not consider during my investigations. For example, my conceptualization of syntactic change abstracted away from the short term effects of priming and the long-term effects of other cognitive mechanisms, was framed in an unnaturally inflexible way in terms of strict competition and replacement, and imagined that changes proceeded linearly from activation to complete penetration in a population.

In many ways then, the mixed-effects models may not capture important real-world aspects of the investigated linguistic changes. I would therefore suggest that one should keep some healthy skepticism towards the findings reported here as the easiest people to fool are ourselves, that one should continue to look for refuting counter-evidence and conduct additional hypothesis tests where possible, and that the model proposed here should serve as a starting point for more particularized investigations, not be viewed as the answer to every word order puzzle posed by Old English conjunctions. However, one should keep in mind that, unless substantial, new Early English texts are unearthed, the statistical estimates presented in this chapter are likely to be as precise as they can possibly become. One should also ensure that one does not place unrealistic demands on the certainty of results based on data from as poorly documented a language as Old English.

Finally, there are various additional questions that arise from this work. Do all conjunctions show the distributional differences between MCs and CCs or are there lexical differences between them? Do conjunct clauses behave more conservatively in poetry than in the prose? Can the decline in Early English argument topicalization be measured more adequately by taking information-structural alignments into account? More generally, how exactly can one design an interface between Old English syntactic rules and their usage in particular discursive contexts? The investigation of this question would have consequences far beyond the specific issue of Early English conjunct clauses. What additional testable predictions follow from the model of C-head conjunctions? For example, it might be possible to trace the development of potential subject topics before negative, subjunctive verbs as items with a high probability to be placed under `C`. Further, the rate of inversion of subject and finite verb should not be significantly different between MCs and CCs if the topic category and the information value of the subject are controlled for. Is it possible to quantify the effect size of the clause type difference between MCs and CCs and show that the strength of the effect can reasonably be assumed to be identical between each of the diverging syntactic patterns? Is there evidence for C-head conjunctions in other early Germanic languages?

Another open question concerns the reasons for the various syntactic developments reported here. However, I fear that an inquiry into this matter would currently be a near-impossible task. The changes are too noisy; potential causal factors are too difficult to measure for each text; the interactions between an individual change and the rest of the grammatical system are too drastic; external confounds from manuscript copying or loan syntax are too strong. If one wants to better understand why a linguistic change occurs, one should use as a case study a change that is easier to measure and documented from start to end in the historical textual record. That is precisely the task I set myself for the next chapter.

# 4

# Approaching the *Why*-Question of Syntactic Change: The Replacement of *then* with *when* in Middle English [1]

*"Prediction is very difficult, especially about the future."*

– Niels Bohr

## 4.1 Introduction

Why language changes arise in the first place is perhaps the most challenging question in historical linguistics. This chapter will attempt to make a modest contribution towards elucidating this complicated issue. The field appears to lack, at the present time, both a definite methodology, i.e., a standardized protocol, for investigating the cause for a particular linguistic change as well as an accepted method of appraisal, i.e. an evaluation metric, to determine if an explanation put forward constitutes a satisfying answer. I will therefore approach the topic with a run-of-the-mill procedure for scientific inquiry - I will observe one particular linguistic change, speculate on possible causes for its occurrence, and derive and experimentally test predictions that follow from the proposed explanation. While this technique may seem trivial and is in fact commonplace in historical linguistics, it still seems worth reminding oneself that the identification of causal factors for a linguistic change is, in principle, no different than any other question about reality and should thus be tackled with the same spirit and principles of the scientific method as for any other empirical problem.

In order to produce satisfying answers to the *why*-question of change, one should find a test case of a linguistic innovation that is relatively simple and clear-cut. Most importantly, it should (i) feature two, easily identifiable variants such that one gradually drives out the other, (ii) involve structures that are quite common so that the rise of the new form can be adequately measured in corpora and (iii) spread from 0-100% of use during the documented history of English. This contrasts sharply, for instance, with some changes analyzed in the previous chapter, which were not straightforward replacements of one syntactic pattern with another, or whose beginnings were missing from the textual record, etc. I believe to have found a change that meets these criteria and should allow a rewarding investigation into its potential causes.

---

[1]Parts of this chapter were presented at the *International Conference of Historical Linguistics 22* (ICHL22) in Naples in July 2015 and at the Ninth Days of Swiss Linguistics (9DSL) in Geneva in June 2016. I would like to thank the participants of these conferences for helpful feedback. The empirical material for this chapter is taken in part from the *Parsed Corpus of Middle English Poetry* (PCMEP). I constructed this database over the course of more than two years from 2012-2015. I am grateful to Benjamin Börschinger and Paola Merlo for useful advice on chunking. I described the construction and purpose of the corpus at the CUSO doctoral workshop *The Media of Literature in the Digital Age* in Geneva in November 2013. The interest of the workshop's attendees greatly encouraged me to proceed with the corpus construction project. The PCMEP was largely compiled during a stay at the University of Pennsylvania in 2014-2015, funded by a Doc.Mobility grant from the Swiss National Science Foundation (P1GEP1_148611), which is hereby gratefully acknowledged. I would like to thank in particular Beatrice Santorini for sharing valuable insights on efficient parsing and many helpful discussion on specific Middle English constructions.

This linguistic innovation concerns a change in the formal realization of adverbial subordinators during the history of English. Old English frequently used the *th*-adverbs *þa* and *þonne* 'then' to introduce temporal adverbial clauses. An example from the West-Saxon translation of the New Testament is presented in (1) below. Here and in the following illustrations, the temporal subordinate clause has been bracketed and the subordinator, sometimes called a conjunction, is shown in boldface.

(1) ac [ **þa** he wæs þæt ger bisceop ] he witgode þæt se hæland sceolde sweltan for ðære þeode
but then he was that year bishop he prophesied that the Savior should die for that people
"But when he was bishop that year, he prophesied that the Savior should die for the people'
(cowsgosp,Jn_[WSCp]:11.51.6768) (c. 990)

From early Middle English times on, the *th*-elements were increasingly replaced by the *wh*-item *when* in this function (e.g., Mitchell 1985: §2775, Declerck 1997: 58-63). Subsequently, there emerged a transitional period with relatively free variation, in which both the conservative and the innovative variant could be encountered. The examples in (2) illustrate this step in the development.

(2) a. Olibrius þe luðere, [ **þa** he þis iherde, ] changede his chere.
Olibrius the wicked when he this heard changed his expression
'Olibirus the wicked changed his expression when he heard this
(CMMARGA,58.55, c. 1225 A.D.)

b. Sothely, þise wordes, [ **when** I here thaym or redis þam, ] stonyes me
truly, these words, when I hear them or read them, stupefy me
'Truly, these words stupefy me when I hear them or read them'
(CMROLLTR,45.918, c. 1345 A.D.)

Eventually, the *th*-forms ceased to be productive and only *when* remained as a grammatical option for the introduction of temporal subordinate clauses. This language stage is exemplified by the sentence in (3), which is a Middle English rendition of the same bible passage as the first example.

(3) but [ **whanne** he was bischop of that ȝeer, ] he prophesiede, that Jhesu was to die for the folc
but when he was bishop of that year he prophesied that Jesus was to die for the folk
'But when he was bishop that year, he prophesied that Jesus would die for the people'
(CMNTEST,11,40J.1143) (c. 1383)

For the purposes of this study, I define as a 'temporal subordinate clause' any finite, adverbial clause introduced by the subordinators *then* or *when*. The specific meaning of such clauses is not important here. For instance, they may also allow a conditional reading. Hence, the label 'temporal' is intended merely as a convenient mnemonic for the investigated structure. The term contrasts with *when*-clauses as complements in indirect questions (*I don't know when I'll be back*) and as generalizing free relatives (e.g., *whenever you are free*, Old/Middle English *when + so*). Thus, the term 'subordinate' refers specifically to non-generalizing adjuncts. Conversely, 'subordinators' or 'subordinating *when*' etc. should be understood to mean the introductory formal markers of temporal subordinate clauses in that sense.

This chapter is structured as follows. I will first describe the corpora used for the retrieval of material for the description of the change from subordinating *then* to *when*. Subsequently, I will use these databases to measure the change and provide some quantitative data on its time course. Third, various possible causes for the change will be outlined. Focusing on one factor in particular, a language-internal trigger in the form of certain word order developments, I will construct a formal LFG model capturing essential features of the relevant structures. The proposed model of the change then allows the derivation of falsifiable hypotheses regarding the purported causal influence of syntactic changes on the emergence of *when* as a subordinator. Fourth, I will proceed to test these hypotheses to determine if the proposed explanation can be supported empirically. Finally, the chapter will conclude with an assessment of the findings and their relevance for a better understanding of linguistic changes in general.

## 4.2 Corpora Used

The empirical material for this chapter is mainly taken from two Middle English corpora. The first is the *Penn-Helsinki Parsed Corpus of Middle English*, second edition (PPCME2) (Kroch and Taylor 2000) (c. 1.4m words), the second the *Parsed Corpus of Middle English Poetry* (Zimmermann 2015) (c. 107 thousand words). The two corpora follow the same annotation guidelines and thus the same search queries can be applied to both of them. In addition, I considered material from Old English texts dated after the year 1000. I will not discuss the Old English texts here, but their names and dates can be found in the description of the YCOE text files in tables 3.1 and 3.2 of the previous chapter. I will use these databases here as the basis for the categorization of syntactic structures into matrix and subordinate clauses, for the statistical counts, and for all example citations.

As for the Old English texts in the last chapter, I assigned to every text file included in the two corpora a definite year of composition based on arguments given in the secondary literature. These dates must again be regarded as approximate and provisional but may be somewhat more certain than the Old English year estimates. I do not have enough space and time to present arguments for the dates for every text file here. However, they do not diverge substantially from the customary dates given to the texts according to the Helsinki system (Kytö 1996), which groups every text into a multi-year time span, most importantly M1 1150-1250, M2 1250-1350, M3 1350-1420 and M4 1420-1500, or according to the Middle English Dictionary system (Lewis et al. 1954–2001), which gives a specific year to every text based on its manuscript witness and, where possible, also an approximate year of composition, e.g., *Treatise on the Astrolabe*, a1450 (c1391) (=manuscript ante 1450, composed c. 1391). My justification for the assignment of a specific year to the poetry texts can be found on the PCMEP corpus homepage at `www.pcmep.net` under the tab 'Text Information.'

The following tables summarize for every text the name of the electronic file, a more customary name for the text, its genre (prose or poetry), and its estimated date of composition. The texts dated before 1200 were already listed in the previous chapter, but are repeated here for convenience. As in the last study, I removed the duplicates and Old English transliterations in the *Lambeth Homilies* and the *Trinity Homilies*.

| Number | File Name | Text Name | Genre / Corpus | Assigned Year |
|---|---|---|---|---|
| 1 | CMPETERB1 | First Chronicle Continuation | Prose / PPCME2 | 1131 |
| 2 | WorcFrag | The First Worcester Fragment | Poetry / PCMEP | 1135 |
| 3 | TheGrave | The Grave | Poetry / PCMEP | 1140 |
| 4 | BodySoul | Body and Soul | Poetry / PCMEP | 1150 |
| 5 | CMPETERB2 | Second Chronicle Continuation | Prose / PPCME2 | 1154 |
| 6 | CMLAMBET | The Lambeth Homilies | Prose / PPCME2 | 1160 |
| 7 | PatNost | Pater Noster | Poetry / PCMEP | 1160 |
| 8 | CMTRINIT | The Trinity Homilies | Prose / PPCME2 | 1160 |
| 9 | CMORM | Ormulum | Poetry / PPCME2 | 1175 |
| 10 | PoemaMorale | Poema Morale | Poetry / PCMEP | 1175 |
| 11 | ProvAlf | The Proverbs of Alfred | Poetry / PCMEP | 1175 |
| 12 | CMVICES1 | Vices and Virtues | Prose / PPCME2 | 1180 |
| 13 | LordOneGod | Lord as Thou art one God | Poetry / PCMEP | 1195 |
| 14 | Ureisun | A Good Orison of Our Lady | Poetry / PCMEP | 1200 |
| 15 | WooingGroup | Texts of the Wooing Group | Prose / PCMEP | 1200 |
| 16 | LAYAMON | Layamon's Brut | Poetry / - (inofficial PPCME2 release) | 1210 |
| 17 | CMANCRIW-1 CMANCRIW-2 | Ancrene Riwle | Prose / PPCME2 | 1220 |
| 18 | Bestiary | Middle English Physiologus | Poetry / PCMEP | 1225 |
| 19 | CMHALI | Hali Meidhad | Prose / PPCME2 | 1225 |
| 20 | CMJULIA | Saint Juliana | Prose / PPCME2 | 1225 |
| 21 | CMKATHE | Saint Katherine | Prose / PPCME2 | 1225 |
| 22 | CMMARGA | Saint Margaret | Prose / PPCME2 | 1225 |
| 23 | CMSAWLES | Sawles Warde | Prose / PPCME2 | 1225 |
| 24 | HarrowHell | The Harrowing of Hell | Poetry / PCMEP | 1230 |
| 25 | LoveRon | The Love Ron | Poetry / PCMEP | 1235 |
| 26 | LittleSerm | A Little Sooth Sermon | Poetry / PCMEP | 1245 |
| 27 | Maximian | Maximian | Poetry / PCMEP | 1245 |
| 28 | ElevenPains | The Eleven Pains of Hell | Poetry / PCMEP | 1250 |
| 29 | OwlNight | The Owl and the Nightingale | Poetry / PCMEP | 1250 |
| 30 | Kentish$_S$ermons | Kentish Sermons | Prose / PPCME2 | 1250 |

Table 4.1: Overview of the PPCME2 and PCMEP text files and their dates of composition, Part 1

| Number | File Name | Text Name | Genre / Corpus | Assigned Year |
|---|---|---|---|---|
| 31 | IacoIose | Jacob and Joseph | Poetry / PCMEP | 1265 |
| 32 | DameSirith | Dame Sirith | Poetry / PCMEP | 1270 |
| 33 | ThruNight | Th Thrush and the Nightignale | Poetry / PCMEP | 1270 |
| 34 | FoxWolf | The Fox and the Wolf | Poetry / PCMEP | 1275 |
| 35 | TreatDreams | A Metrical Treatise on Dreams | Poetry / PCMEP | 1285 |
| 36 | Havelok | Havelok the Dane | Poetry / PCMEP | 1290 |
| 37 | Husbandman | Song of the Husbandman | Poetry / PCMEP | 1297 |
| 38 | Cokaygne | The Land of Cockayne | Poetry / PCMEP | 1300 |
| 39 | Fridesw | The Legend of St. Frideswide | Poetry / PCMEP | 1305 |
| 40 | DavyDreams | Adam Davy's Five Dreams | Poetry / PCMEP | 1310 |
| 41 | Marina | The Life of St. Marina | Poetry / PCMEP | 1320 |
| 42 | OrisFiveJoys | Orison of the Five Joys | Poetry / PCMEP | 1320 |
| 43 | Simonie | The Simonie | Poetry / PCMEP | 1325 |
| 44 | CMAYENBI | Ayenbite of Inwyt | Prose / PPCME2 | 1340 |
| 45 | CMROLLEP | Richard Rolle's Epistles | Prose / PPCME2 | 1345 |
| 46 | CMROLLTR | Richard Rolle's Treatises | Prose / PPCME2 | 1345 |
| 47 | DisMaryCross | Dispute between Mary and the Cross | Poetry / PCMEP | 1350 |
| 48 | CMEARLPS | The Earliest Prose Psalter | Prose / PPCME2 | 1350 |
| 49 | HowHearMass | How to Hear Mass | Poetry / PCMEP | 1355 |
| 50 | CMGAYTRY | Dan Gaytryge's Sermon | Prose / PPCME2 | 1357 |
| 51 | CMEDTHOR | The Mirror of Edmund (Thornton) | Prose / PPCME2 | 1360 |
| 52 | WynWas | Winner and Waster | Poetry / PCMEP | 1360 |
| 53 | CMMANDEV | The Travels of Sir Mandeville | Prose / PPCME2 | 1371 |
| 54 | CMEDVERN | The Mirror of Edmund (Vernon) | Prose / PPCME2 | 1375 |
| 55 | CMAELR3 | Rievaulx De Institutione (Vernon) | Prose / PPCME2 | 1375 |
| 56 | CMBOETH | Chaucer's Boethius | Prose / PPCME2 | 1380 |
| 57 | CMNTEST | Wycliffe's New Testament | Prose / PPCME2 | 1383 |
| 58 | CMOTEST | Wycliffe's Old Testament | Prose / PPCME2 | 1383 |
| 59 | CMWYCSER | Wycliffite Sermons | Prose / PPCME2 | 1383 |
| 60 | 1385-CHAU-TRO | Chaucer's Troilus and Cressida (excerpt) | Poetry / - | 1385 |
| 61 | CMPOLYCH | John Trevisia's Polychronicon | Prose / PPCME2 | 1387 |
| 62 | CMPURVEY | Purvey's Prologue to Wycliffe | Prose / PPCME2 | 1388 |
| 63 | CMJULNOR | Julian of Norwich's Revelations | Prose / PPCME2 | 1390 |
| 64 | 1390-CHAU-SQT | Chaucer's Squire Tale (excerpt) | Poetry / - | 1390 |
| 65 | CMASTRO | Chaucer's Treatise on the Astrolabe | Prose / PPCME2 | 1391 |
| 66 | CMEQUATO | Equatorie of the Planets | Prose / PPCME2 | 1392 |
| 67 | CMCTPARS | Chaucer's Parson's Tale | Prose / PPCME2 | 1395 |
| 68 | CMCTMELI | Chaucer's Tale of Melibee | Prose / PPCME2 | 1395 |
| 69 | CMCLOUD | The Cloud of Unknowing | Prose / PPCME2 | 1395 |
| 70 | SirCleges | Sir Cleges | Poetry / PCMEP | 1395 |
| 71 | CMHILTON | Hilton's Eight Chapters on Perfection | Prose / PPCME2 | 1396 |
| 72 | BirdFoFe | The Bird with Four Feathers | Poetry / PCMEP | 1400 |
| 73 | CMVICES4 | Book of Vices and Virtues | Prose / PPCME2 | 1400 |
| 74 | CMBRUT3 | The Chronicles of England | Prose / PPCME2 | 1400 |
| 75 | LetterCupid | The Letter of Cupid | Poetry / PCMEP | 1402 |
| 76 | CMTHORN | Liber de Diversis Medicinis | Prose / PPCME2 | 1410 |
| 77 | CMMIRK | Mirk's Festial | Prose / PPCME2 | 1410 |
| 78 | CMBENRUL | Northern Rule St. Benet | Prose / PPCME2 | 1415 |
| 79 | CMROYAL | Sermons from Ms. Royal | Prose / PPCME2 | 1425 |
| 80 | CMHORSES | A Treatise on Horses | Prose / PPCME2 | 1425 |
| 81 | CMAELR4 | Rievaulx De Institutione (Bodley) | Prose / PPCME2 | 1430 |
| 82 | CMKEMPE | The Book of Margery Kempe | Prose / PPCME2 | 1435 |
| 83 | CMEDMUND | The Life of St. Edmund | Prose / PPCME2 | 1438 |
| 84 | CMCAPSER | Capgrave's Sermon | Prose / PPCME2 | 1452 |
| 85 | CMCAPCHR | Capgrave's Chronicle | Prose / PPCME2 | 1460 |
| 86 | CMGREGOR | Gregory's Chronicle | Prose / PPCME2 | 1467 |
| 87 | CMMALORY | Malory's Morte Darthur | Prose / PPCME2 | 1470 |
| 88 | CMREYNES | Reynes' Commonplace Book | Prose / PPCME2 | 1475 |
| 89 | CMREYNAR | Caxton's Reynard$_t$he$_F$ox | Prose / PPCME2 | 1481 |
| 90 | CMFITZJA | Fitzjames' Sermo die Lune | Prose / PPCME2 | 1495 |
| 91 | CMINNOCE | In Die Innocencium | Prose / PPCME2 | 1497 |
| 92 | CMSIEGE | The Siege of Jerusalem | Prose / PPCME2 | 1500 |

Table 4.2: Overview of the PPCME2 and PCMEP text files and their dates of composition, Part 2

The data base comprises a total of 92 Middle English + 39 Old English = 131 text files. They include c. 1.6 million words of running texts in 117,503 sentence tokens. Figure 4.1 represents the relative size and date of origin for every text file. The red dots depict the poetic records, the blue dots prose texts. As the graph shows, there is a substantial gap in the transmission of Middle English prose texts in Helsinki period M2, i.e., between c. 1250 and 1350. The prose lacuna appers to be even greater than the poor text transmission after the Norman Conquest in 1066. The gap is often used as a convenient way to distinguish between early (c. 1100-1350) and late Middle English (c. 1350-1500), although the precise division is not standardized. As the red dots show, the verse texts can bridge the gap in the prose of the late 13$^{th}$ and early 14$^{th}$ centuries. They are therefore crucial witnesses of English during this era. There are many more poetic documents from this time not included here, which could further close the gap in prose in future studies. All in all, the poetry has about one tenth of the size of the prose in my data set.

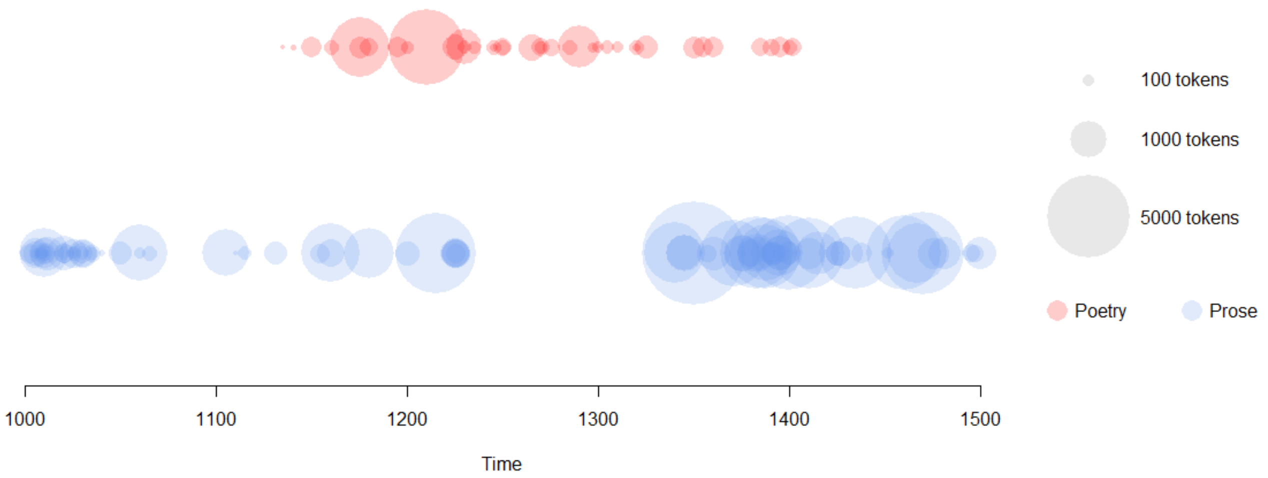



Figure 4.1: Temporal distribution and size of poetry and prose considered in this study

Unlike the predominantly West-Saxon Old English texts, the Middle English witnesses come from a wide range of scribal dialects. Moreover, the preponderance of specific dialects shifts during the course of the Middle English period. The East-Midlands dialect becomes dominant while the West-Midlands and Southern dialects become less important from the late 14$^{th}$ century on. Northern sources enter the textual record only from the middle of the 14$^{th}$ century on whereas Kentish dies out as a distinct written form at that time. Furthermore, the scribal dialects of a manuscript witness are often demonstrably different from the dialect of their underlying original exemplar, indicating substantial dialect mixing. The fact that the different dialects are unevenly distributed across the Middle English era and often mixed makes it very difficult to control for dialect differences systematically in statistical models of morpho-syntactic changes. The reason is that the variation one encounters cannot unambiguously be attributed to dialect differences but could also be due to uncertain dialect attributions, diachronic developments or even authors' individual stylistic choices. Dialect differences will therefore not be investigated in great detail in this chapter.

I will now proceed to measure the variation between *when* and *then* as temporal subordinators in the texts listed above.

## 4.3 Measuring the Rise of *when* as a Subordinator

In this section, I will describe the measurement of the replacement of *then* with *when* as temporal subordinators in Middle English. I will first describe how I collected the data. Subsequently, I will discuss and evaluate the results of this investigation.

### 4.3.1 Data Collection

I collected all Middle English temporal subordinate clauses that were introduced by the subordinator *then* or *when*. Adjunct clauses of the relevant kind are annotated as PPs with a P-head and an adverbial clause complement, `CP-ADV`. Structures were not considered if the adverbial clause complement contained an empty *wh*-element in its specifier, i.e., when the subordinate clause resembled a comparative clause. Where one subordinator introduced coordinated embedded clauses, only the first conjunct was considered. The `P`-head contains the crucial information about the formal realization of the subordinator. The difference between the *th*- and *wh*-based subordinator was operationalized in terms of a set of strings contained in a definition file. There is substantial spelling variation in Middle English and hence the sets of spelling variants were relatively long. The lists in (4 ) show some examples of strings meant to instantiate *then* and *when* respectively.

(4) **Example strings for the retrieval of *then* and *when***

a. `THEN: +ta|+Ta|+da|+Da|+tonne|+Tonne|+donne|+Denne|+denne|+Donne| ...`

b. `WHEN: Hw*n*|hw*n|W*n*|w*n*|H*n*|h*n*|Q*n*|q*n|Uu*n*|uu*n*|$Hw*n*| ...`

It is worth noting that the conservative variant *then* actually consists of two different forms.The first Middle English *th*-subordinator is *þo*, which derives from Old English *þa*. The other *th*-item is *þan*, which goes back to Old English *þonne*. Scholars have proposed three kinds of differences between these two forms. First, the two items may have distinct functional or discourse-pragmatic properties (Wårvik 1995). Second, there may be dialectal differences. For example, in Southern and West Midlands texts of the thirteenth century, variants of *þan* are uncommon (e.g., *King Horn*, D-Version of *Poema Morale*, Katherine Group, Robert of Gloucester's *Chronicle*) (Kivimaa 1966). In general, however, dialect differences are difficult to distinguish from random text effects. Lastly, there appear to be clear semantic differences. These have been pointed out for Old English by Yamakawa (1969). He claims that Old English *þonne* points to an unspecific time in the future or a generality whereas Old English *þa* points back to a definite, past event.

> Loosely speaking, we can say that *þonne* and *þa* are synonymous with each other [...]. With more semantic precision, however, it must be asserted that there is a fairly clear distinction in general between these two particles. *þonne* as a conjunction [...] is used when the time of an action or occurrence is indefinite and general or it is to be habitually repeated, and is usually found with a [...] verb in the present tense [...]. On the other hand, *þa* as a conjunction [...] is used when the narrator is going to describe a definite action or occurrence confined to a particular point of time, and is most commonly found with a [...] verb in the past tense (Yamakawa 1969: 11)

A brief investigation into Middle English texts shows that the same effect obtains there. I subdivided all hits I retrieved for the conservative variant into present tense and past tense clauses. Modal auxiliaries, `MD`, are not annotated for tense in the corpora and were therefore all grouped into the former tense category. I then divided the data further into instances of structures headed by *þan*, which were operationalized as forms of *then* that contained the grapheme *n*, and adjunct clauses headed by *þo*, which were defined as instances of *then* that did not contain the grapheme *n*. The resulting 2x2 table is presented in table 4.3 below.

| Tense | *þo* | *þan* |
|---|---|---|
| present | 23 | 247 |
| past | 231 | 45 |

Table 4.3: Tense in temporal subordinate clauses by subordinator form (Middle English texts only)

While the results may not be absolutely accurate owing to wrong annotations, a relatively coarse definition of the difference between the two *th*-based subordinators and noise introduced by modals, the overall trend is very clear - *þo* tends to occur with past tense verbs whereas *þan* is usually found with present tense verbs ($\chi^2$=307.0, df = 1, $p$<0.001). This finding confirms Yamakawa's Old English semantic generalization for Middle English texts as well. Two typical examples are presented in (5) (see also example (36) in the preceding chapter).

(5) a. **Present tense - *þan***

and [$_{CP}$ **þan** hie fulle beð. ] hie secheð to þe fule floddri.
and then they full are they seek to the foul mire

'And when they are full, they desire the foul mire [of drunkenness]' (CMTRINIT,37.511)

b. **Past tense - *þo***

[$_{CP}$ **tho** I escaped from hym ] I loste myn one ere
then I escaped from him I lost my one ear

'When I escaped from him, I lost one of my ears' (CMREYNAR,52.309)

Even though dialect and semantic differences are discernable between *þan* and *þo*, I will group both kinds of *th*-based temporal subordinators together in the remainder of this chapter. The reason is that both items undergo the same change - they both compete with, and are eventually replaced by, *when*.

The subordinator *when* is extremely rare in my Old English data. Temporal subordinate clauses are usually introduced by *þa*, *þonne* 'then' instead. I found only three relevant instances of *when*, annotated as free temporal relatives, `CP-FRL-TMP*`. These examples are presented in (6).

(6) **Old English temporal subordinate clauses introduced by *when***

a. ... þæt he wyle abidan, [$_{\text{CP}}$ **hwænne** he hire eað gewrecan muge ]
... that they will abide when they theirs easily avenge might
'...so that they will wait [instead] when they might easily avenge their [sense of injury]'
(coalcuin,Alc_[Warn_35]:215.154)

b. minre sawle þyrste to þan lyfigende Gode,
my soul thirst.subjunctive to the living God
[$_{\text{CP}}$ **hwænne** ic cume & me æteowige beforen Godes ansene ]
when I come and reflexive appear before God's sight
'May my soul be thirsty for the living God when I come and appear before the face of God'
(coalcuin,Alc_[Warn_35]:292.213)

c. ... [$_{\text{CP}}$ **hwænne** he us to cyme ] we witan mid gewisse þæt hit þærto nealæcð georne.
... when he us to come we know with certainty that it thereto approaches eagerly
'... when he [=the judge/God] comes to us, we will know with certainty that it [=the appointed day/the end of the world] is very near'
(cowulf,WHom_2:62.43)

In addition, my Old English data set includes two examples of temporal subordinate clauses introduced by *when* modifying a nominal antecedent, (7)

(7) **Old English *when*-clauses modifying a temporal antecedent**

a. Gesete me andagan, [$_{\text{CP}}$ **hwænne** ðu wille þæt ic for þe gebidde ]
set me time when you want that I for you pray
'Set me a time when you want me to pray for you.' (cootest,Exod:8.9.2629)

b. & seotte þa dæi [$_{\text{CP}}$ **hwonne** man scolde þæt mynstre gehalegon ]
and set the day when one should that minster hallow
'And he fixed the day when the minster should be consecrated'
(ChronE_[Plummer]:656.22.407)

I conducted two straightforward search queries for the variants of the dependent variable, i.e., one for *then* the other for *when*. Every example was annotated for its year of composition, as shown in tables 4.1 and 4.2 for Middle English and the texts dated after the year 1000 in the text tables of the preceding chapter for Old English, its source text, and the genre of its source text (prose vs. poetry). In total, I conducted 2 (*then* vs. *when*) * 2 (two periods, Old and Middle English) = 4 search queries.[2]

### 4.3.2 Results

I retrieved 776 (Old English) + 546 (Middle English) = 1,322 examples of *then*-clauses and 5 (Old English) + 3,970 (Middle English) = 3,975 instances of *when*-clauses, i.e.. a total of 5,297 examples. Next, I fitted a mixed-effects model to the temporal subordination data. It predicts the realization of the subordinator as a *wh*-item from time and genre and includes a random text effect. The resultant model is given in table 4.4.

[2] The definition of *then* and *when*, the CorpusSearch query and output files and an excel file summarizing the data can be found on the Dissertation DVD under `Chapter 4 - WhenThen/01 Rise in When`.

```
formula = When ~  Year + Genre + (1 | Text),
        family = binomial, data = ThenWhen

Fixed effects:
                            Estimate      Std.Error z-value p
Intercept                   -39.963       3.0834    -12.961 <0.001***
Year                        0.033084      0.002474  13.374  <0.001***
Genre(Prose→Poetry)         -1.0820       0.5206    -2.078  0.0377*

Random effect:
Text, N=124
Variance of random intercepts:  3.598

Null deviance:          5952 on 5296 degrees of freedom
Residual deviance:      1255 on 5293 degrees of freedom
AIC: 1163
```

Table 4.4: Logistic Regression Model 1: mixed-effects model for the rise in *when* as a subordinator

The model estimates an increase in *when*-subordination of 0.033 log-odds per year. At this rate of change, it would take 277 years for the innovative subordinator to rise from 1% to 99% of use. The time variable is highly significant in this model. As the genre category is changed from prose to poetry, the log-odds of *when* decrease slightly by -1.1, i.e., the innovative variant is somewhat less likely to appear in poetic documents than in prose texts. The genre variable is a significant predictor at the 5% level in this model. This shows that the change in subordination would conform to the general belief that poetry, especially medieval poetry, tends to be linguistically more conservative than prose (e.g., Nielsen 2003, much more nuanced Fulk 1992: §§43–45). However, the effect size is relatively weak. The subordination data comes from a total of 124 texts. That means that 7 texts do not contain any relevant instances of *then* or *when*. The random variability among these texts is very substantial as shown by a relatively large variance parameter of 3.6.

The change displays a constant rate effect with respect to genre. In other words, prose and poetry seem to innovate *when* as a temporal subordinator at the same speed. This is shown by the fact that the addition of the interaction term between genre and time does not reduce deviance sufficiently to warrant its inclusion in the model ($\chi^2$=0.095, df=1, $p$=0.758). If one presupposes, as I think one reasonably can, that a constant rate effect should be found in the present context, its actual detection would suggest the following two corollaries. First, the dates assigned to the prose and poetry texts seem to be relatively accurate. Otherwise, an erroneous text chronology would likely have distorted the time course of the change and resulted in more divergent time coefficients for the two genres. Secondly, and more importantly, the finding shows that the change in subordination affects poetic and prose texts alike. The poetry texts can therefore fruitfully be used to measure the change and fill the gap in prose texts between c. 1250–1350.

The model performs very well, especially given that it represents a change from medieval times. It fits significantly better than a null model (where the null model is an intercept only model) (Likelihood Ratio Test, $\chi^2$= 4856.8, df=3, $p$<0.001). It fits the data adequately overall (Pseudo-$R^2_{marginal}$= 0.770, Pseudo-$R^2_{conditional}$=0.890). The model has substantial classificatory power over the baseline (C-index= 0.990, classification accuracy =96.0% vs. baseline: 75.0%).

The graph in figure 4.2 illustrates the data I collected for the change from *then* to *when*. Every text is represented with one data point. Prose documents are shown in blue, verse texts in red. The red dots tend to fall slightly below the blue dots, indicating the slightly greater conservativeness of the poetry. The graph also shows that the red dots can successfully close the prose gap in period M2, c. 1250–1350, which underlines the usefulness of the PCMEP text files. The figure includes two independent regression lines for prose in blue and for poetry in red as well as their 95%-confidence intervals. The presence of a constant rate effect is suggested by the fact that two independent regressions for the two genres return very similar year coefficients of 0.033 for prose and 0.032 for poetry respectively.

All in all, the rise in *wh*-based subordination seems to be very clear-cut, especially for a change measured in medieval texts. It is simple in that there is competition between two variants that are easily identifiable by their form, *then* and *when*, and that can therefore straightforwardly be measured in corpora. Furthermore, the replacement is completely attested in the history of English from its earliest beginnings to its complete penetration in the population. Documents before c. 1125 almost categorically use a form of *then* for temporal subordination whereas texts after c. 1375 almost universally show a form of *when*. Virtually all free variation occurs within the narrow 250-year transitional period in between. Thus, this innovation should provide a good testing ground for posing and approaching the why-question of linguistic change.

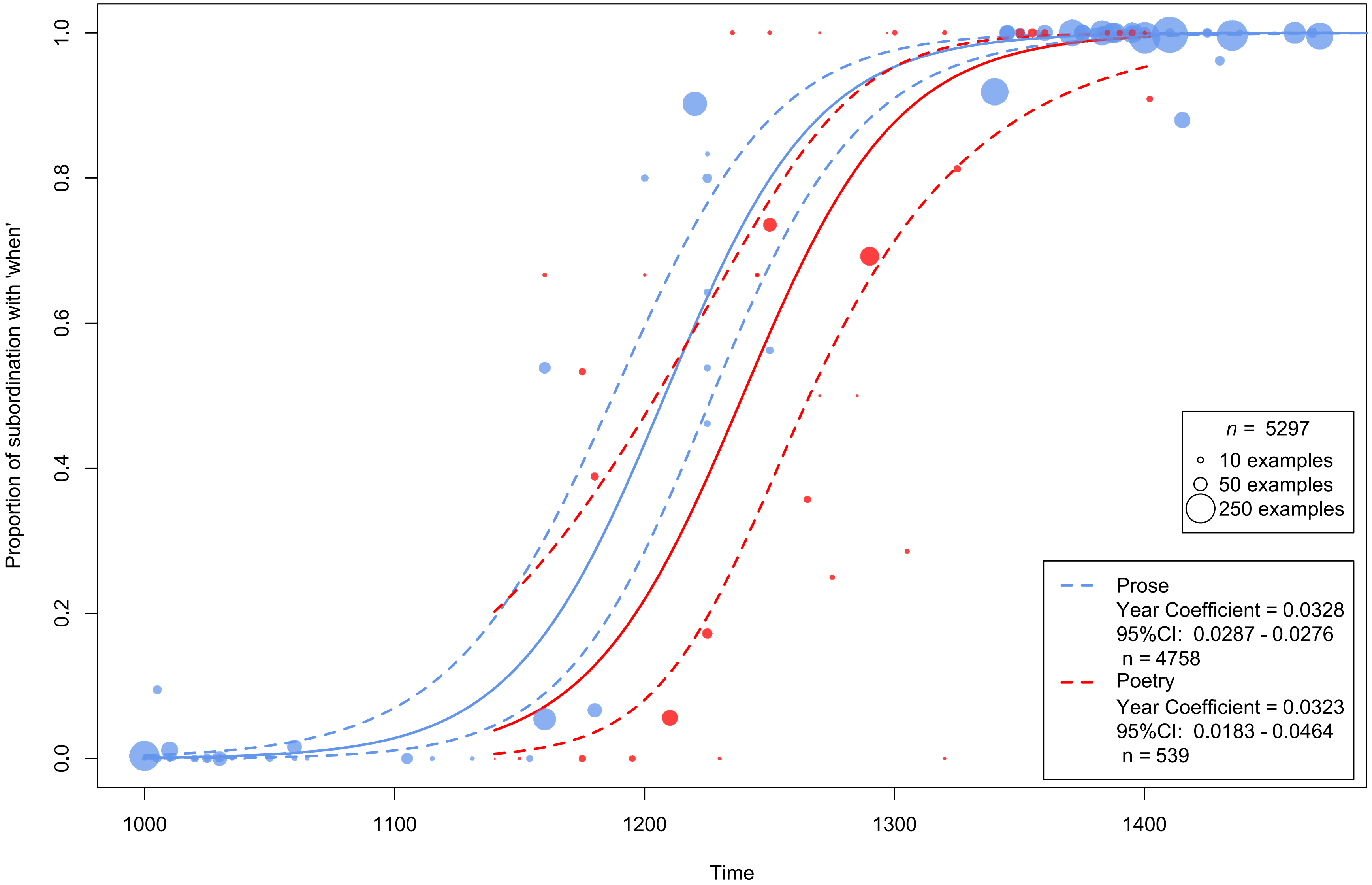



Figure 4.2: Illustration of the rise in temporal subordination with *when*

I will conclude this section by presenting a few more examples of the variation between subordination with *then* and *when*. Specifically, the following sentence pairs from poetic documents of the transitional period illustrate the presence of both the conservative and the innovative form within a single text, hence, one may reason, the representation of both rules for the generation of the conservative and the innovative variant within the mind of a single author or speaker.

(8) *The Rhymed Pater Noster (Lambeth Homily no. [6])*, c. 1160

a. we nabben wil to sunegen.
we not-have will to sin

[$_{CP}$ **þenne** ure unwines us munegen ]
then our enemies us tempt

'We have no will to sin when our enemies tempt us' (PatNost,72.63.144)

b. to gode solf we us wreið;
to God self we us accuse

[$_{CP}$ **hwenne** we þos word seggeð ]
when we these words say

'We denounce ourselves to God himself when we say these words' (PatNost,93.65.182)

(9) *The Bestiary*, c. 1225

a. we ben siker dere, So ðis wirm in winter is, [$_{CP}$ **ðan** ge ne tileð nummore ]
we are safe there, as this worm in winter is, then you not till nomore

'We will be safe there, just as this insect is in winter, when you do not till anymore' (Bestiary,143.9.290.[Ant_Significance])

b. [$_{CP}$ **wanne** he is ikindled ] Stille lið ðe leun,
when he is whelped still lies the lion

'When he is first born, the lion lies quietly' (Bestiary,9.1.17.[Lion_Nature])

(10) *The Owl and the Nightingale*, c. 1250

a. þan gode ich fulste to longinge,
the good I help in longing

vor [$_{CP}$ **þan** him longeþ, ] ich him singe:
for then him longs I him sing

'I help the good man in longing, for when he feels desire, I sing to him' (OwlNight,76.890.503)

b. an prostes upe londe singeþ
and priests upon land sing

[$_{CP}$ **wane** þe liȝt of daie springeþ]
when the light of day springs

'Country priests sing when dawn breaks' (OwlNight,62.734.423)

(11) *Havelok the Dane*, c. 1290

a. [$_{CP}$ **þan** he was ded, ] þere micte men se
then he was dead, there might men see

þe meste sorwe that micte be
the most sorrow that might be

'When he was dead, one could see, the greatest sorrow that could ever be' (Havelok,8.233.107)

b. [$_{CP}$ **Hwan** he wore come, ] sket was þe erl yare,
when he was come, quickly was the earl ready,

Ageynes denshe men to fare,
against Danish men to go

'When he had arrived, the earl was ready right away to advance against the Danish' (Havelok,73.2575.1188)

Having presented quantitative data on the change from *then* to *when*, I will now consider several possible reasons for the execution of the change.

## 4.4 Why was *then* Replaced by *when*?

In this section I will attempt to offer a possible explanation for the change from subordinating *then* to *when*. My approach to investigating why a particular linguistic change occurred consists of (a) speculating on possible factors that could have contributed to the change and (b) deriving testable hypotheses that could lend plausibility to them. Such factors should not be thought of as definite determinants of the change, but rather as salient aspects of the development, i.e., as some of the few identifiable components of a much larger mental, social and historical system that is so complex as to not permit, at least at the present time, comprehensive and realistic modeling. In particular, I will surmise that word order changes led to a linguistic environment in which *when* would receive an advantage over *then* and construct a formal grammar model that embodies the relevant constructions. I will then derive from it a few predictions that should be true if the proposed explanation plays a role in the change.

### 4.4.1 List of Possible Influencing Factors

Several factors can be isolated as potential causes for the development of a linguistic system that would favor *when* over *then* for temporal subordination. I will mention three of the most noticeable ones.

(1) First, Middle English is characterized by extensive borrowing from Norman French. It is commonly believed that about 10,000 lexemes were brought into English from that language (e.g., Kastovsky 2006). Furthermore, French may have asserted substantial influence on the realization of Middle English syntactic constructions in such a way that grammatical patterns of the donor language were translated by lexical material available in the source language. For example, French may have provided a template for the formation of periphrastic adjectival comparison (e.g., González-Diaz 2008) or non-finite perfect *have* after modals (e.g., Ingham 2010*a*), etc. Thus, it does not seem unreasonable that the intensive language contact with Norman French would also have resulted in calquing the Old French *wh*-item *quant* (or perhaps Latin *quando*) as *when* to introduce temporal subordinate clauses.

I will illustrate the possible influence of French on the change with a citation from a rare Anglo-Norman – Middle-English parallel text. William Twiti, huntsman to King Edward II (1307-1327), drafted a prose treatise on the art of hunting in Norman French. John Gifford made a Middle English translation of the text in the early 15$^{\text{th}}$ century. Example (12a) presents a *quant*-clause in the French original. The sentence in (12b) is its English parallel correspondence and involves a *when*-clause.

(12) a. *Anglo-Norman French*

E [$_{\text{CP}}$ **quant** il est de quatre annees ] il doit partyr hors de la soundre par age.
and when he is of four years he must part out of the sounder by age.

'And when he [=boar] is four years old, he ought to depart out of the sounder because of his age.'
William Twiti's *Le Art de Venerie*, edition: Dryden (1843: 2), c. 1320

b. *Middle English*

and [$_{\text{CP}}$ **when** they be of .iiij. yere age ] they shall departe fro the sounder for age
and when they be of four year age they shall depart from the sounder for age

'And when they [=boars] are four years old, they will depart from the sounder because of their age.'
John Gifford's translation, *Le Venery de Twety*, edition: Wright and Halliwell (1845: 151), c. 1410

The French borrowing scenario is also supported by the fact that some Middle English texts, perhaps especially those from the North, spell *wh*-words with *qu*, reminiscent of the spelling of French interrogative function words. An example from *Cursor Mundi* is shown below.

(13) Forþi blisce I þat paramour
therefore bless I that loved-one

[$_{\text{CP}}$ **Quen** I haue nede me dos socure, ]
when I have need me does assistance

'Therefore I bless that loved one [who] when I have need, helps me'
*The (Northern) Cursor Mundi*, edition Morris (1874-1893: 12, lines 69-70), c. 1300

(2) Next, one could invoke the principle of analogy to explain the change. Middle English re-purposes *wh*-items for a wide range of grammatical, especially relative, structures.

For example, nominal relative clauses were originally formed with demonstratives functioning as relative pronouns or the indeclinable relative particle *þe*, as in (14a), but Middle English begins to use *wh*-items for that purpose, as in (14b). The two examples are an Old and a Middle English translation respectively of an identical passage from the Gospel of John.

(14) a. **Relative clause formed with indeclinable particle *þe***

æfter me cymð [$_{\mathsf{DP}}$ wer [$_{\mathsf{CP}}$ **þe** me beforan geworden wæs ]]
after me comes man that me before become was

'A man comes after me who was made before me' (cowsgosp,Jn_[WSCp]:1.30.5791)

b. **<u>Middle English innovation</u>: subject relativized by *which***

Aftir me is comun [$_{\mathsf{DP}}$ a man, [$_{\mathsf{CP}}$ **which** was maad bifor me ]]
after me is come a man which was made before me

'A man has come after me who was made before me' (CMNTEST,1,20J.63)

Likewise, locative relative clauses were typically expressed by the *th*-element *there*, as exemplified in (15a). Middle English texts recruit the *wh*-item *where* to realize such relatives, as in (15b). The two examples are taken from the same text illustrating competition between the two forms by a single speaker.

(15) a. **Locative relative introduced by *there***

in [$_{\mathsf{DP}}$ the same place [$_{\mathsf{CP}}$ **there** the grete batayle was, ]] ys grete tresoure hydde in the erthe
in the same place there the great battle was is great treasure hidden in the earth

'In the same place where the great battle was, a great treasure was hidden in the earth' (CMMALORY,30.947)

b. **<u>Middle English innovation</u>: place relativized by *where***

I com but late oute of [$_{\mathsf{CP}}$ the Waste Foreyste [$_{\mathsf{CP}}$ **where** I founde the Rede Knyght ]]
I came but late out of the waste forest where I found the red knight

'I came only late out of the desolate forest where I had found the red knight' (CMMALORY,667.4880)

The same is true for all kinds of free relatives, which used to relativize grammatical functions with a *th*-based element in Old English but have come to be expressed with *wh*-items from Middle English times on. For example, a free relative object was typically introduced by the complementizer *that*, (16a), but now requires the *wh*-pronoun *what*, (16b). The two examples below are taken from the same text, illustrating, once again, grammar competition within a single individual. I represent free relativization with a vacuous DP-projection introducing a zero antecedent.

(16) a. **Argument free relative introduced by *that***

Ich have y-doo [$_{\mathsf{DP}}$ [$_{\mathsf{CP}}$ **þat** y schulde ]]
I have done that I should

'I have done what I should' (CMPOLYCH,VIII,111.3726)

b. **<u>Middle English innovation</u>: argument free relative formed with *what***

here after y schal doo [$_{\mathsf{DP}}$ [$_{\mathsf{CP}}$ **what** me nedeþ ]]
here after I shal doo what me needs

'Hereafter, I shall do what I need' (CMPOLYCH,VIII,111.3727)

In this connection, one can also mention the grammaticalization of the Old English relative clause construction *the while(s) that*, (17a), to the independent Middle English subordinator *while*, (17b). The second example is the earliest citation of *while* as a independent subordinator of word class *conj.* (conjunction) in the Oxford English Dictionary. It is possible that this development was buttressed by the general rise of *wh*-items illustrated in the previous examples on account of the fact that both *while* and the *wh*-elements feature initial *wh*-graphemes and are hence formally similar.

(17) a. **Old English *the while that* construction**

and he wunode þa swa on his broðor hirede [$_{DP}$ **þa hwile** [$_{CP}$ **ðe** he leofode ]]
and he lived then so in his brother's court the while that he lived

'And he lived thus in his brother's court while he was alive'
(Chronicle2:ChronC_[Rositzke]:1041.3.1801)

b. **Middle English innovation: introduction of subordinator *while***

& ðat lastede þa xix wintre [$_{CP}$ **wile** Stephne was king ]
and that lasted the 19 years while Stephen was king

'This lasted the nineteen years while Stephen was king' (CMPETERB2,56.452)

Thus, the change from subordinating *then* to *when* dovetails with a more general spread of *wh*-items. It is therefore conceivable that the change occurred largely as a consequence of analogical reasoning, i.e., as a diffusion of a subordinating pattern from one *wh*-item to the next, whose ultimate trigger could in principle be completely unrelated to determinants directly affecting the rise of subordinating *when*.

(3) Finally, it is possible that the change resulted from a loss of certain disambiguating word order patterns. This is the option that I will focus on in this chapter.

The Old English operator adverbs *þa*, *þonne* attracted the verb in their use as main clause operator adverbs 'then' to the high verb position C, leading to verb-second patterns. This aspect of Old English syntax was formalized in section 3.3.1.4 of the previous chapter. The operator feature on *þa* and *þonne* is gradually lost so that these items come to be parsed as ordinary sentential modifiers that are followed by canonical 'subject - verb' order instead. I specified this change as the loss of V-to-C structures after operator adverbs in the same place. I repeat an example of verb-second order after *then* for convenience in (18) (=(40) of chapter 3).

(18) **Verb-second after operator adverb *then***

[$_{CP}$ **þa** [$_{C'}$ *aras* [$_{IP}$ he ]]]
then arose he

'Then he arose' (cowsgosp,Mk_[WSCp]:2.14.2301)

In contrast, pre-historic Old English *þa* and *þonne* were probably followed quite systematically by verb-final order in their use as a temporal subordinator 'when,' as suggested by the fact that verb-final patterns decline during Early English and were thus probably more common at some earlier point (e.g., see Hypothesis 5 on the decline of verb-final structure in section 3.4.3.5 of the previous chapter). An example of a verb-final temporal *then*-clause is presented in (19). There might thus have been complementary distribution between the interpretation of *then* as an operator adverbs in context *A*, i.e., before verb-second structures, and as a subordinator in context *B*, i.e., before verb-final patterns.

(19) **Verb-final after subordinator *then***

ac sume dæge on ærnemergen [$_{CP}$ **þa** [$_{IP}$ he of slæpe *awoc*, ]] he abræc into ðam bure
but some day on early-morning then he of sleep awoke he broke into the bower

'But some day, in the morning, when he awoke from sleep, he broke into her bedchamber'
(coapollo,ApT:1.10.9)

By the time of recorded Old English, subordinating *þa* and *þonne* could quite commonly occur with verb-medial structure in the embedded clause. Example (20) shows a verb-medial temporal *then*-clause. The syntactic conditioning of the interpretation of *then* might thus better be thought of as a contrast between an adverbial reading in context *A*, i.e., before verb-second structures, and a subordinator use in context *not A* or 'elsewhere,' i.e., not before a verb-second structure (for the view that verb-final order is not a perspicuous marker of subordination, see Stockwell and Minkova 1990: 509).

(20) **Verb-medial after subordinator *then***

... gelice þam þe Iudeas didon [$_{CP}$ **þa** [$_{IP}$ hi *mængdon* eced and geallan togædere ]]
... like to-that that Jews did then they mixed vinegar and gall together

'... similar to what the Jews did when they mixed vinegar and bile together'
(cocanedgD,WCan_1.1.1_[Fowler]:39.48)

A summary of the Old English conditioning factors for the interpretation of *then* is shown in (21) below, where the forward slash means 'in the context of.'

(21)

| | | | |
|---|---|---|---|
| THEN | → | main clause operator adverb | / __ verb-second |
| | → | embedded clause temporal subordinator | / ∼ (elsewhere) |

The conditioning factors shown above weaken over time. Specifically, *then* loses its special status as an operator adverb, and with it, its ability to attract the finite verb to the high verb position leading to verb-second patterns. The categorical constraint "verb second order occurs if and only if *then* is an adverb" becomes a merely probabilistic constraint, "verb second order occurs if *then* is an adverb, but other alignments are also, at first, rarely, later sometimes, eventually most of the time, compatible with its interpretation as an adverb." Hence, the syntactic conditions no longer consistently distinguish between the two functions of *then* - canonical 'subject - verb' word order can now occur after *then* both in its use as a main clause adverb and as a dependent clause subordinator. Middle English speakers, as modern readers of Middle English texts, have to rely exclusively on the linguistic context to decide which grammatical role was intended. The following example illustrates a case of clause-initial *then* followed by 'subject - verb' order that seems to be interpretable to the same degree of plausibility as an adverb or as a subordinator.

(22) **Tho** the screwe was overcome; Sori he was and wo.
then/when the villain was overcome, sorry he was and woe
(i) 'Then the villain was overcome. He was sorry and miserable.' (adverb)
(ii) 'When the villain was overcome, he was sorry and miserable.' (subordinator)
(Fridesw,43.55)

Thus, the loss of conditioning word order patterns results in ambiguous interpretations of *then*. The creation of this kind of ambiguity could then plausibly have contributed to the emergence of *when* as a designated subordinator. This idea has been entertained by other scholars before. For instance, Yamakawa gives the following, somewhat obscure, formulation of what seems to be a similar idea about the cause for the change from subordinating *then* to *when*.

> "Although the delicate shift in syntactic nature from paratactic determinative to hypotactic connective, often perceived in instances of *þanne* or *þo*, offers us suggestive data for considering the historical movement in the sphere concerned, it implies, at the same time, a weak point in functional distinctness that *þanne* or *þo* has when used as a subordinate or relative conjunction. This seems to be an internal motive that caused *þanne* or *þo* to be replaced by *hwanne*." (Yamakawa 1969: 31)

Thus, there also seems to be some scholarly precedence for the idea that changes in certain word order patterns might be regarded as an antecedent cause for the rise of subordinating *when*. However, it still remains unclear how exactly one should conceive of the causal role of the syntactic developments on the change in question. I can think of two different conceptualizations.

(i) **Evolutionary perspective**. A designated subordinator that can unequivocally indicate subordinate clause readings, such as *when*, will have a greater fitness or an advantage over a more opaque competitor that cannot, such as *then*. It may be possible to quantify the advantage of one subordinator over the other as the number of cases that it alone can unambiguously parse (similar to proposals in Yang 2002). The exact nature of factors that may confer an advantage of one form over another are not fully understood at the present time. However, it is clear that the assessment of a form's fitness is relative to the grammar environment in which it is used. Here, subordinating *when* would only be superior to subordinating *then* because the linguistic context that used to bestow recognizability or copying fidelity onto *then* is disappearing. In my opinion, a conceptualization of the change along the lines of such evolutionary dynamics is the most appealing and holds the greatest promise for future derivations of testable predictions.

(ii) **Functional perspective**. Alternatively, one could conceptualize the change as a kind of syntactic push chain shift. If the output of a sound change, e.g., Grimm's Law changing the Proto-Germanic voiced plosive /b/ to the voiceless plosive /p/, encroaches on the phonetic space of an existing sound, e.g., /p/, the linguistic system often moves the straitened sound away, e.g., shifting the voiceless plosive /p/ to the voiceless fricative /f/, thus preserving the distinctness of the two phonological categories, e.g., /b∼p/ vs. /p∼f/. There are substantial treatises on the cognitive motivations for such processes (e.g., Martinet 1952 on the tendency of phonemes to maximize the security space around them, Labov's (1994: 119) 'Principle of Contrastiveness', also Labov 2010). In like manner, if the result of a syntactic change, e.g., the replacement of verb-second order after adverbial *then* with 'subject - verb' order after adverbial *then*, leads to formal similarity with another syntactic pattern, e.g., 'subject - verb' order after subordinator *then*, the linguistic system may attempt to alter the expression of the infringed syntagma, e.g., replacing subordinating *then* with subordinating *when*, thus preventing the merger of two syntactic categories, e.g., *then*-initial matrix clauses vs. temporal subordinate clauses. The cognitive functions offered as explanations for phonological chain shifts may carry over to their syntactic counterparts straightforwardly. In my opinion, such a functional viewpoint is only partially convincing because it comes dangerously close to providing an untestable, teleological formulation of the change.

I would like to comment in a little more detail on the difference between evolutionary and functional perspectives on linguistic changes. Evolutionary explanations invoke differential degrees to which competing linguistic forms are adapted to a grammar system or ancillary systems in the explanans. For the case at hand, one could reason in evolutionary terms along the following lines. The innovative form *when* should be recognized as a subordinator almost without exception in a subordinating context. It would thus frequently be involved in the analysis of a linguistic structure and therefore become more entrenched in the mind over time. The conservative form *then*, on the other hand, would not be unambiguously recognized as a subordinator whenever it occurs in a subordinating context. It would thus be activated less frequently in parsing and therefore become less strongly represented in a population of grammars as time goes by. In contrast, functional explanations refer to a certain purposiveness of the mechanisms referenced in the explanans. In the present case, typical functional description could read as follows. One could say that speakers might have recruited *when* <u>in order to</u> preserve the contrast between adverb-initial and temporal subordinate clauses, or that the subordinator *then* was abandoned <u>for the purpose</u> of preventing the collapse of those two constructions, or that the employment of the *wh*-item *when* constitutes a kind of repair <u>strategy</u>, etc. In this connection, I am reminded of the twofold wording of genetic changes in Richard Dawkins' famous book *The Selfish Gene* - one formulation in "respectable" mechanistic terms of natural selection and fitness, the other formulation involving a more "sloppy" short-hand metaphor of genes as conscious agents with aims.[3] In the same way, the evolutionary and functional perspectives on innovations in a grammar could be regarded as merely different ways of conceptualizing the same empirical phenomenon. However, " respectable" explanations formulated in evolutionary terms should still be regarded as more appropriate than "sloppy" formulations within functional frameworks. The former can potentially account for the illusion of design and purpose of a linguistic change. The latter merely enshrines this illusion in a large number of disjointed principles and tendencies.

In summary, I listed three potential causes for the development of *when* as a subordinator - (1) Norman-French calquing processes, (2) analogical extensions originating in the spread of other subordinating *wh*-items and (3) the loss of verb-second word order conditioning the interpretation of *then* as an adverb. This list is not comprehensive but merely comprises the factors whose causal role in the change seemed most convincing to me by intuition.

In principle, all three factors could be investigated simultaneously. For this endeavor to work, one would have to collect data in which each of the above conditions appear independently of each other and in which cells of the experimental set-up do not remain empty. However, the obtainment of such data is not feasible in the present case. The reason is that the incidence of French loans, *wh*-based subordination, and 'subject - verb' order after *then* all increase in parallel at the same time so that the values of these factors will be correlated with each other. Such multicollinearity between factors affecting a linguistic change will make it impossible for a multiple regression to yield valid estimates of the effect of each individual predictor. To give just one example, one would want to find authors whose language remained relatively unaffected by Norman French influence, a variable that could be operationalized as the number of French loan words, for instance, but who nevertheless use a fair number of *wh*-subordinators, such as relative *which*, *where what* etc. Only such texts could allow for the estimation of the effect of analogical extension independently of the effect of French borrowings. Unfortunately, such texts do not exist, let alone in sufficient numbers. These reflections highlight (i) the enormous difficulty in disentangling the effects of multiple predictors on a linguistic outcome variable outside of controlled experiments in a laboratory and (ii) the fact that a satisfactory analysis of possible causal factors for a linguistic change are entirely contingent on the quality of the available data.

I will focus on only one of the developments identified as potentially causative - the changes in the syntactic word order patterns that used to regulate the interpretation of *then*. Admittedly my findings will be relatively uncertain seeing that they ignore the influence of other, potentially antithetical factors. Nevertheless, a closer inspection of the role of the syntactic developments in isolation can at least offer some support for the assumption that they exert a causal influence on the change at hand.

[3] "If we allow ourselves the licence of talking about genes as if they had conscious aims, always reassuring ourselves that we could translate our **sloppy** language back into **respectable** terms if we wanted to, we can ask the question, what is a single selfish gene trying to do?" Dawkins, Richard (1976) *The Selfish Gene*, chapter 6: Genesmanship. Oxford: Oxford University Press, page 88, emphasis mine.

### 4.4.2 A Model of Old and Middle English *when* and *then*

In this section, I will construct an LFG toy grammar for the rise of *when* as a temporal subordinator in Middle English. The model will be able to analyze *then*-initial matrix clauses with and without inversion, temporal subordinate *when*- and *then*-clauses, as well as a few related constructions.

#### 4.4.2.1 Main Clause Adverb *then*

As a starting point, I will repeat my model assumptions for the analysis of *then* as a main clause adverb from the preceding chapter. The conservative Old English operator adverbs *þa* and *þonne* 'then' contain a presentation link feature. This is exemplified by the lexicon entry for *þonne* in (23) (=(37) of chapter 3).

(23) *þonne* ADV (↑ PRED)= 'THEN'
(↑ PRES-LINK)= +

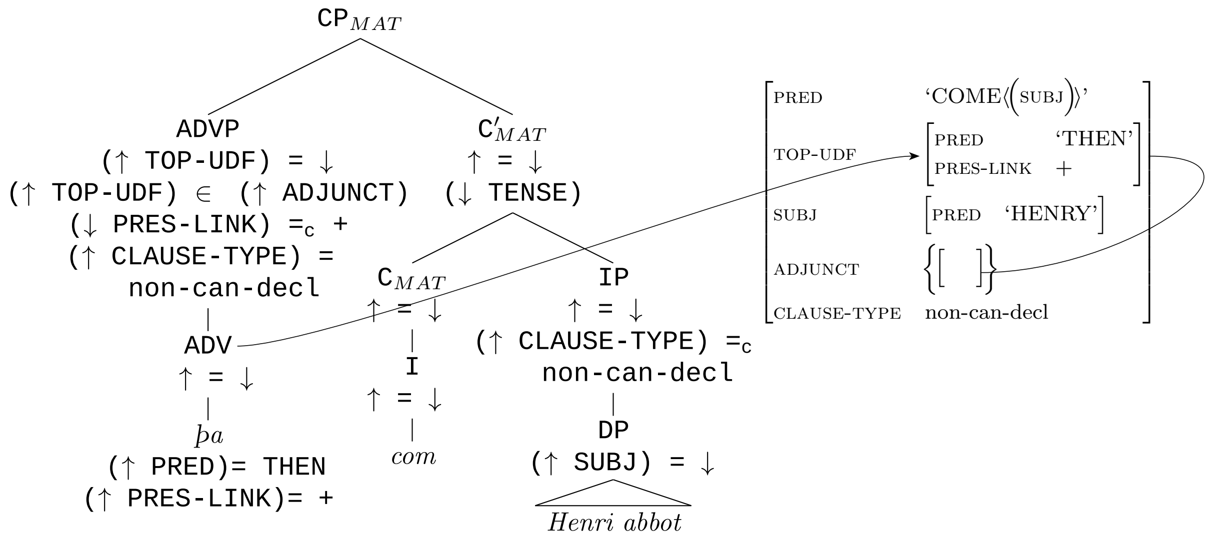
Operator adverbs are parsed with a rule specifically associated with such items. This rule screens an adverb in the specifier of $\mathtt{CP}_{MAT}$ for the presence of a presentational link feature. If it is encountered, the adverb is licensed in this position and a non-canonical declarative clause type feature is introduced triggering high verb placement under C. The operator adverb rule is repeated below in (24) (=(38) of chapter 3).

(24) $\mathtt{CP}_{MAT} \rightarrow$ ADVP $\mathtt{C'}_{MAT}$

| | ADVP | $\mathtt{C'}_{MAT}$ |
|---|---|---|
| $\mathtt{CP}_{MAT} \rightarrow$ | (↑ TOP-UDF) = ↓ | ↑ = ↓ |
| | (↑ TOP-UDF) ∈ (↑ ADJUNCT) | (↓ TENSE) |
| | (↓ PRES-LINK) $=_c$ + | |
| | (↑ CLAUSE-TYPE) = non-can-decl | |

An example of a Middle English clause involving the operator adverb *then* is shown below in (25). The operator adverb is placed in the specifier of CP and is followed by verb second order. I assume that full subjects almost universally occur in the specifier of IP in Middle English and cannot remain lower in the structure anymore as in Old English. This idealization is incorrect for early Middle English, but no important consequences hinge on this assumption. I sketch a rudimentary f-structure for the parse. The analysis functions exactly as described before and I will therefore not comment on it again.

(25) þa com Henri abbot
then came Henry abbot
'Then the abbot Henry came' (not: # 'when the abbot Henry came')
(CMPETERB2,54.370)




My model can generate verb-second order after *then* only by rule (24). The verb immediately following *then* is thus a necessary and sufficient condition for its interpretation as a main clause adverb. This view is widely shared but not universally accepted. For example, Andrew (1940, see also Campbell 1970: 95) believes that verb-second order is compatible with subordinate clause readings. His argument is the following. He investigates passages ...

> "... from works which are translations from the Latin: [...]
>
> *Oros. 156.29.* Ða ascedan hiene his þegnas ... Ða ondwyrde he.
> *234.21.* Ða bæd he þæt mon þone triumphan him ongean brohte. Ða sende mon ...
>
> So, Sweet, making two pairs of demonstrative sentences. But in each passage the Latin begins with a temporal clause 'cum a sociis increperatur' and 'cum nuncios de victoria misisset', and again there can be little doubt that we should render 'when his thanes asked him' and 'when he demanded a triumph'. [...] Such passages establish quite clearly the 'þa com he' type of subordinate clause as a genuine OE idiom." (Andrew 1940: §§11–12)

Andrew seems to have a point. On closer inspection, I would concur with him that there may be instances of *then*-initial clauses with verb-second order that are quite naturally interpreted as dependent clauses. The following example illustrates.

(26) Ða het se papa hine to biscope gehalgian & hine on Breotone sende.
then ordered the pope him to bishop consecrate and him to Britain sent

**Ða com** he ærest upp in Westseaxum & heo þær hæðne gemette,
then came he first up to West-Saxons and them there heathens met

þa ðuhte him nyttre & betre þæt he ðær Godes word bodade & lærde.
then seemed him more-useful and better that he there God's word preached and taught

'Then the pope had him ordained bishop and sent him to Britain.
(i) # Then / at that point he first came to the West-Saxons and met them there as heathens;
(ii) When he first came to the West-Saxons and met them there as heathens,
then it seemed better to him that he should preach God's word'
(cobede,Bede_3:5.166.27.1609-1613)

Here, it seems unsuitable to interpret *þa com he* as 'then he came.' The interpretation of *then* as a marker of successive temporal order, 'next, afterwards,' seems inappropriate because the phrase describes the first item in a list of events as indicated by *ærest* 'first.' Its interpretation as a temporal deictic, 'at that point,' is hard to accommodate also because the time reference of 'coming to the West-Saxons' does not equate to the reference point of the preceding clause, 'being consecrated bishop' or 'being sent to Britain.' In contrast, the interpretation of *then* as a temporal subordinator, 'when,' makes good sense, as shown in the translation, and corresponds to the Latin original (Andrew 1940: §11) as well as the punctuation used in the edition.

Despite the fact that there may be a chance that *then* immediately followed by the verb may not necessarily establish its reading as a main clause adverb, I will not adapt my model assumptions. The reasons for this are as follows. (i) The implementation of a model that can handle paratactic embedding, i.e., high verb placement under C with simultaneous subordination, would require an amount of revision incommensurate with the goals of this chapter. (ii) It is quite clear that the majority of verb-second *then*-clauses are interpreted as main clauses. The relation between word order and interpretation assumed here is therefore valid at least as a strong tendency. (iii) It is possible that Old English translations freely render Latin subordinate clauses as main clauses. In this case, Anglo-Saxon translators may have approached their Latin sources more creatively than Andrew gives them credit for. All in all, then, I believe that it is a defensible idealization to model a categorical association between verb-second order after *then* and its usage as a temporal adverb (see also Mitchell 1985: §2539).

Now, at some point, Early English develops innovative ordinary adverbs that are formally identical to the operator adverbs *þa* and *þonne* 'then' but that do not contain a presentation link feature. A lexicon entry for such an adverb is shown in (27) (= (43) of chapter 3).

(27) *than* ADV (↑ PRED)= 'THEN'

Ordinary adverbial forms of *then* are parsed with a general rule for all kinds of sentential modifiers. This rule disallows an adverb featuring salient syntactic features, such as a *wh*-feature, a presentational link feature, or a negative polarity feature, and, if such a feature is in fact absent, adjoins the adverb to the category $\mathrm{CP}_{MAT}$.[4] The rule does not introduce a clause type feature and, all things being equal, will therefore create a structure with the verb inside the IP, commonly Middle English 'subject - verb' order. The sentential adverb rule is repeated below in (28) (=(41) of chapter 3, with the ban on the presence of a negative polarity feature added).

[4] Appendix A discusses the Modern English ban on adjoining to CP an element that includes a negative polarity feature in connection with negative PPs to rule out structures such as **Under no circumstances, Mary would misbehave.*

(28) $\texttt{CP}_{MAT} \rightarrow$ $\quad \underset{\substack{\downarrow\, \in\, (\uparrow\, \{\texttt{XCOMP|COMP}\}^{*}\ \texttt{ADJUNCT}) \\ (\downarrow\ \texttt{ADV-TYPE}) \neq \texttt{wh} \\ (\downarrow\ \texttt{POL}) \neq \texttt{neg} \\ \neg(\downarrow\ \texttt{PRES-LINK}) = +}}{\texttt{ADVP}} \quad \underset{\uparrow\, =\, \downarrow}{\texttt{CP}_{MAT}}$

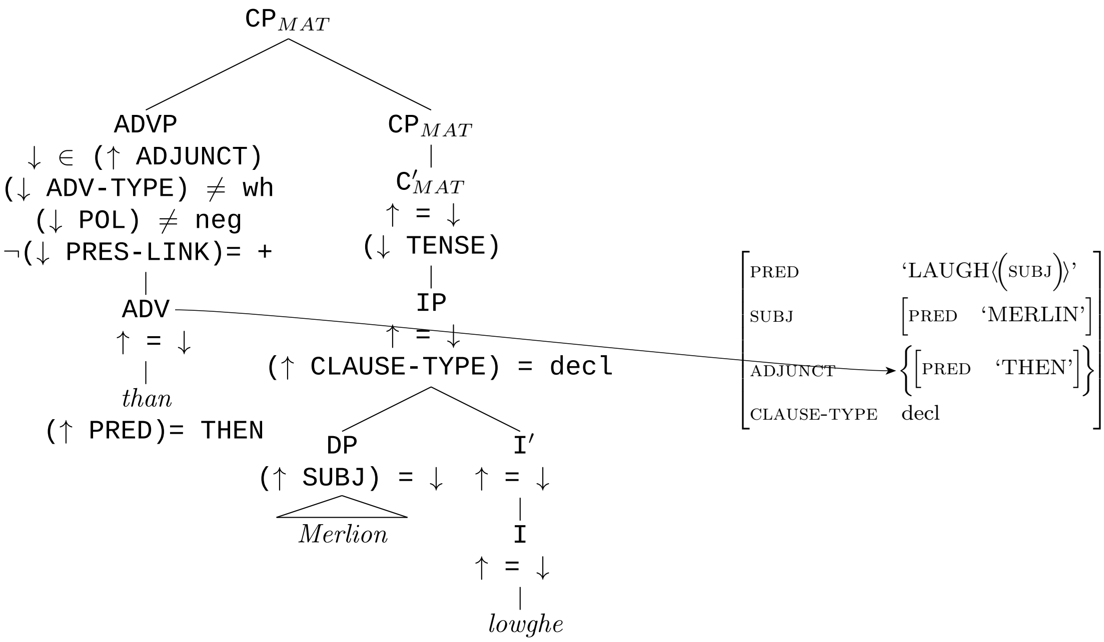
Hence, my model assumes that just as Modern English $\texttt{CP}_{MAT}$ cannot be modified by *wh*-items, (29a), or negative adverbs, (30a), but by non-*wh*-items, (29b), or positive adverbs, (30b), so, too, its Early English ancestor did not accept adjunction of the operator adverb *then*, but allowed such a configuration for the innovative non-operator *then*.

(29) a. * [$_{\texttt{CP}}$ Why, [$_{\texttt{CP}}$ Mary kissed John. ]]
b. [$_{\texttt{CP}}$ Therefore, [$_{\texttt{CP}}$ Mary kissed John. ]]

(30) a. * [$_{\texttt{CP}}$ Never, [$_{\texttt{CP}}$ Mary kissed John. ]]
b. [$_{\texttt{CP}}$ Sometimes, [$_{\texttt{CP}}$ Mary kissed John. ]]

An example of Middle English *then* as an innovative sentential modifier is given in (31). The sentential adverb is adjoined to CP and is followed by 'subject - verb' word order. The subject is assumed to be necessarily placed in the specifier of IP. I show only a rudimentary f-structure, including the most essential syntactic features. The parse works according to the same principles as described earlier and I will therefore not repeat a detailed explanation here.

(31) Than Merlion lowghe
then Merlin laughed
'Then, Merlin laughed' (context makes it clear that not: # 'when Merlin laughed')
(CMMALORY,70.2375)




Conservative operator adverb *then* as in (23) competes with and is eventually replaced by innovative sentential modifier *then* as in (27). This, in effect, models the decline in verb-second after adverbial *then*. The change was illustrated on page 140 in terms of the general model of the implementation of a linguistic innovation in a group of speakers from the introductory chapter.

#### 4.4.2.2 Subordinating *then* and *when*

I will now turn to the subordinator use of *then* and *when*. I will analyze temporal subordinate clauses introduced by these items as free relatives. There are several reasons for this decision. (i) There seems to be a consensus among experts that Modern English temporal subordinate *when*-clauses have the internal structures of free relatives. One can thus simply appeal to authority (e.g., Geis 1970, Grimshaw 1977, Bresnan and Grimshaw 1978, Declerck 1997, Haegeman 2010). (ii) The temporal clause's time variable modifying the main clause event can be bound directly within the subordinate clause or within a finite clause more deeply embedded inside of it. Ambiguities between local and long-distance extraction of this kind are typical of relative structures (cf. *the man who Mary promised to marry* → *Mary promised her father to marry* vs. *Mary promised to marry Bill*). A typical example from Modern English is shown in (32).

(32) I saw Mary in New York **when** [$_{\mathtt{IP}}$ she claimed [$_{\mathtt{CP}}$ that [$_{\mathtt{IP}}$ she would leave. ]]]
(i) high construal: at the time that she made the claim
(ii) low construal: at the exact time of her presumed departure
(from: Haegeman 2010: 635, her ex. (21))

It is difficult to demonstrate this effect for Old and Middle English. However, it seems unwarranted to assume that the potential for long-distance extraction of the temporal subordinator should not have existed in those language stages too. In addition, I would like to offer the following two Early English attestations as plausible instances of long-distance extraction of an initial temporal subordinator from its local predicate's complement clause. The time point variable is realized by *then* in the Old English sentence in (33a) and by *when* in the Middle English clause in (33b).

(33) a. ... he of deaðe aras,
... he of death arose

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **þa** | [$_{\mathtt{IP}}$ he cydde | [$_{\mathtt{CP}}$ þæt | [$_{\mathtt{IP}}$ he ær | mihte | ful | eaðe | deað | forbugan | gyf | he | swa | wolde ]]] |
| then | he announced | that | he earlier | might | full | easily | death | escape | if | he | so | wanted |

' ... [that] he arose from death when he earlier announced that he would very easily escape from death if he so desired' (cowulf,WHom_6:182.363)
(i) probably not high construal: # at the time that he made the announcement
(ii) probably low construal: at the exact time of his announced escape from death

b.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| sum | vnseli | haueð | **hwenne** | [$_{\mathtt{IP}}$ ha seide | [$_{\mathtt{CP}}$ [$_{\mathtt{IP}}$ ha schriue hire. ]]] | ischriuen | hire | alto | wunder. |
| some | un-souly | had | when | she said | she shrived her, | shriven | her | al-to | wonder |

'[Only] a wretched woman has, when she said she shrove herself, [actually] shriven herself wondrously' (CMANCRIW-1,II.56.537)
(i) probably not high construal: # at the time that she made the assertion
(ii) probably low construal: at the exact time of her supposed self-confession

(iii) The treatment of *when*-clauses as free relatives unifies them with other *wh*-based relative constructions that emerge during Middle English. The analysis thus concretizes the hypothesized role of analogical reasoning as a cause for the change from *then* to *when*.

Accepting these arguments as valid, I will proceed to implement temporal subordinate clauses as free relatives in the following way. The outer category of free temporal relatives are modeled as adverb phrases, ADVP. The introduction of such clauses can thus be accomplished with the same rule regulating the incorporation of regular adverbs shown in (28). However, temporal subordinate clauses can also appear in clause-final position, and hence the model requires an additional rule for the adjunction of the ADVP to the right of CP. This additional adjunction rule is presented in (34). It is identical to the one in (28) in all respects except for the relative order between main clause and adjunct. As before, the adjunction rule allows the adverb phrase to be interpreted local to the main clause or within a more deeply embedded main clause complement. With respect to temporal clauses, the latter option can handle, for instances, cases such as *When hell freezes over, Mary said that she would go out with him _.*

(34)
$$
\mathtt{CP}_{MAT} \rightarrow \quad
\begin{array}{c}
\mathtt{CP}_{MAT} \\
\uparrow = \downarrow
\end{array}
\quad
\begin{array}{l}
\qquad\qquad \mathtt{ADVP} \\
\downarrow\, \in (\uparrow \{\mathtt{COMP|XCOMP}\}^{*}\ \mathtt{ADJUNCT}) \\
\quad (\downarrow \mathtt{ADV\text{-}TYPE}) \neq \mathtt{wh} \\
\quad (\downarrow \mathtt{POL}) \neq \mathtt{neg} \\
\quad \neg(\downarrow \mathtt{PRES\text{-}LINK}) = +
\end{array}
$$

The adverb phrase framing free relatives does not occur with an overt head. Instead, the mere application of its re-write rule will introduce a PRED feature as the head of the adjunct function. I will value this feature as PRO, which indicates some deictic time point variable and could be paraphrased as 'at that time.' The adverb phrase rewrites directly to a relative clause. I indicate its category as a CP and assign the specific clause type 'relative,' $\texttt{CP}_{REL}$. The domain of the relative clause maps onto a function that I call 'relative clause modifier,' abbreviated RELMOD. It is essentially an adjunct formally realized as a relative clause. The rule is shown below in (35).

(35) ADVP → $\texttt{CP}_{REL}$
(↑ PRED) = 'PRO'
↓ ∈ (↑ RELMOD)

I will now discuss some basic aspects of English relative clauses, $\texttt{CP}_{REL}$. My discussion is inspired by the account offered in Dalrymple (2001: 400-5), but deviates from it substantially in some details. The following four parsing aspects are shared by all English relative clauses. (i) The specifier of $\texttt{CP}_{REL}$ can be instantiated by a wide range of syntactic categories, such as DPs (*the woman who came*), PPs (*the woman in whom I trust*), APs (*a deed proud of which I should not be*), etc. (ii) The phrase in the specifier of $\texttt{CP}_{REL}$ will be interpreted as an underspecified discourse function that must be unified with a local or long-distance related function internal to the relative clause. The initial function appears to be some kind of topic as the conditions on its extraction resemble those of ordinary topics (Dalrymple 2001: 403-4). I will therefore refer to the constituent in the specfier of $\texttt{CP}_{REL}$ as a relative topic. (iii) The phrase in $\texttt{CP}_{REL}$'s specifier establishes a co-reference relationship with the antecedent of the relative clause. It must therefore contain an anaphoric item, often called a relative pronoun. Hence, I follow a mediated, binding analysis of relative clauses here. I will not be concerned with the exact binding principles governing an appropriate relation between antecedent and relative topic. The anaphoric items are formally marked in some way. For instance, in Modern English, they are realized as *wh*-elements (*who*, *whom*, *which* etc.). The *wh*-item can be encountered on the relative topic directly (*a book [which] I like*), or on its possessor (*a book [whose cover] I like*), or on one of its semantically restricted complements (*a book [for which] I waited*), etc. (iv) The relative clause must be finite.

The rule in (36) presents an abstract template for English relative clauses that incorporates the four points just discussed. (i) The specifier of $\texttt{CP}_{REL}$ is indicated as a meta-variable, XP, which stands for any kind of adequate syntactic category in this context. (ii) It is interpreted as a relative topic, which I label a relative clause underspecified discourse function, REL-UDF, and must unify along a path with some function internal to the relative clause, say its object, one of the members of its adjunct set, the object of its complement and so on and so forth. (iii) The relative topic must require at some depth the presence of some attribute that should be valued in some way to permit its interpretation as a relative pronoun. For example, in Modern English the relative topic, or optionally its possessor or a restricted object function, must contain a pronoun type feature that is valued as *wh*, (↓ (POSS|$\texttt{OBJ}_{th}$) PRON-TYPE)$=_c$ wh. (iv) The finiteness requirement is implemented by endowing $\texttt{C'}_{REL}$ with a constraint, identical to the one for matrix clauses, that calls for the introduction of a tense feature.

| (36) $\texttt{CP}_{REL}$ → | XP<br>(↑ REL-UDF) = ↓<br>(↑ REL-UDF) = some path to<br>a function within the relative clause<br>(↓ REL-UDF some path to an attribute)<br>$=_c$ some value | $\texttt{C'}_{REL}$<br>(↓ TENSE) |
|---|---|---|

For the kind of free relatives under investigation here, a head of the relative clause does not need to be projected Instead, $\texttt{C'}_{REL}$ can directly re-write to IP. As in matrix clauses, the absence of subject verb inversion will introduce a declarative clause type feature at the IP-level. The resultant rule is shown in (37). The remainder of the relative clause IP will be generated with rules already formalized earlier.

(37) $\texttt{C'}_{REL}$ → IP
↑ = ↓
(↑ CLAUSE-TYPE) = decl

The final step in my analysis consists of phrase structure rules specific to relative clauses introduced by the subordinators *then* or *when*. These rules are essentially a concrete realization of the abstract relative clause template discussed for (36) above. It must have been at this stage of the sentence generation algorithm that a small change took place in Middle English. It mutated the conservative rule incorporating *then* as a relative topic so that there emerged an innovative alternative rule permitting the *wh*-item *when* to be used in this function instead.

I will first discuss the conservative, Old English variant of the rule. It can be formalized as follows. The specifier of $\texttt{CP}_{REL}$ is filled with an adverb phrase. It will be interpreted as a relative topic, `REL-UDF`, and must be unified with a member of the local or long-distance related adjunct set, `ADJUNCT`. Now, the relative topic must be headed by a form of the lexeme *then.* There are various possibilities to restrict the applicability of the relativization rule accordingly. First, one could leave the relation between the relative topic and its head unspecified and let their meaning regulate their appearance. Hence, one would trust that principles of semantic composition might rule out ungrammatical structures in which just about any adverb would be placed in the relative topic position (*the time* { *when* / **now* / **soon* / **yesterday* } *I woke up*). However, since I do not implement the semantics of temporal subordinate clauses here, the model would contain quite a large explanatory gap. Secondly, the relative clause rule could require the presence of a formal feature inside the relative topic that *then* is endowed with. Such a feature could refer to *then*'s status as a temporal deictic, its shape as a *th*-item or perhaps the operator feature, `PRES-LINK`, introduced earlier to force verb-second order in main clauses. The problem with this approach is that the rule would be engaged in a wider context than just temporal subordination and therefore make diachronic claims that are more specific than I would like to entertain. For example, if the conservative relativization rule targets all *th*-based items (*then*, the Old English demonstratives, *there*, *thence*, *thither* etc.), relative constructions with those items should all decline at the same rate of change. Similarly, if the conservative rule is sensitive to the `PRES-LINK` feature, one would expect that the decline in subordinating *then* occurs at the same rate of change as the decline in verb-second order after operator adverb *then.* Finally, one might restrict the conservative relative clause rule specifically to the item *then.* This is the option I decided to implement here. Hence, I placed a constraint on the adverb phrase placed in the specifier of $\texttt{CP}_{REL}$ that checks for the valuation of its `PRED`-feature as an instance of *then.* The lexicalized constraints on the rule may seem somewhat unusual and inelegant, but closely corresponds to the scope of my theoretical assertions.

(38) $\texttt{CP}_{REL} \rightarrow$

| | |
|---|---|
| ADVP | $\texttt{C'}_{REL}$ |
| $(\uparrow$ REL-UDF$) = \downarrow$ | $\uparrow = \downarrow$ |
| $(\uparrow$ REL-UDF$) \in (\uparrow \{$COMP\|XCOMP$\}^*$ ADJUNCT$)$ | $(\downarrow$ TENSE$)$ |
| $(\downarrow$ PRED$) =_c$ THEN | |

The innovative, Middle English variant of the temporal relativization rule is largely identical to its conservative competitor. The only difference is that the relative topic must be headed by a form of *when.* A lexicon entry for the word *when* is shown in (39). It is identical in structure to the *wh*-adverb *why* shown in (47) of chapter 2. The item could thus be employed without any further modifications in the generation of direct questions using the rules created in the same place (e.g., *When did you come?*).

(39) *when* ADV $(\uparrow$ PRED$)=$ 'WHEN'
$(\uparrow$ ADV-TYPE$)=$ wh

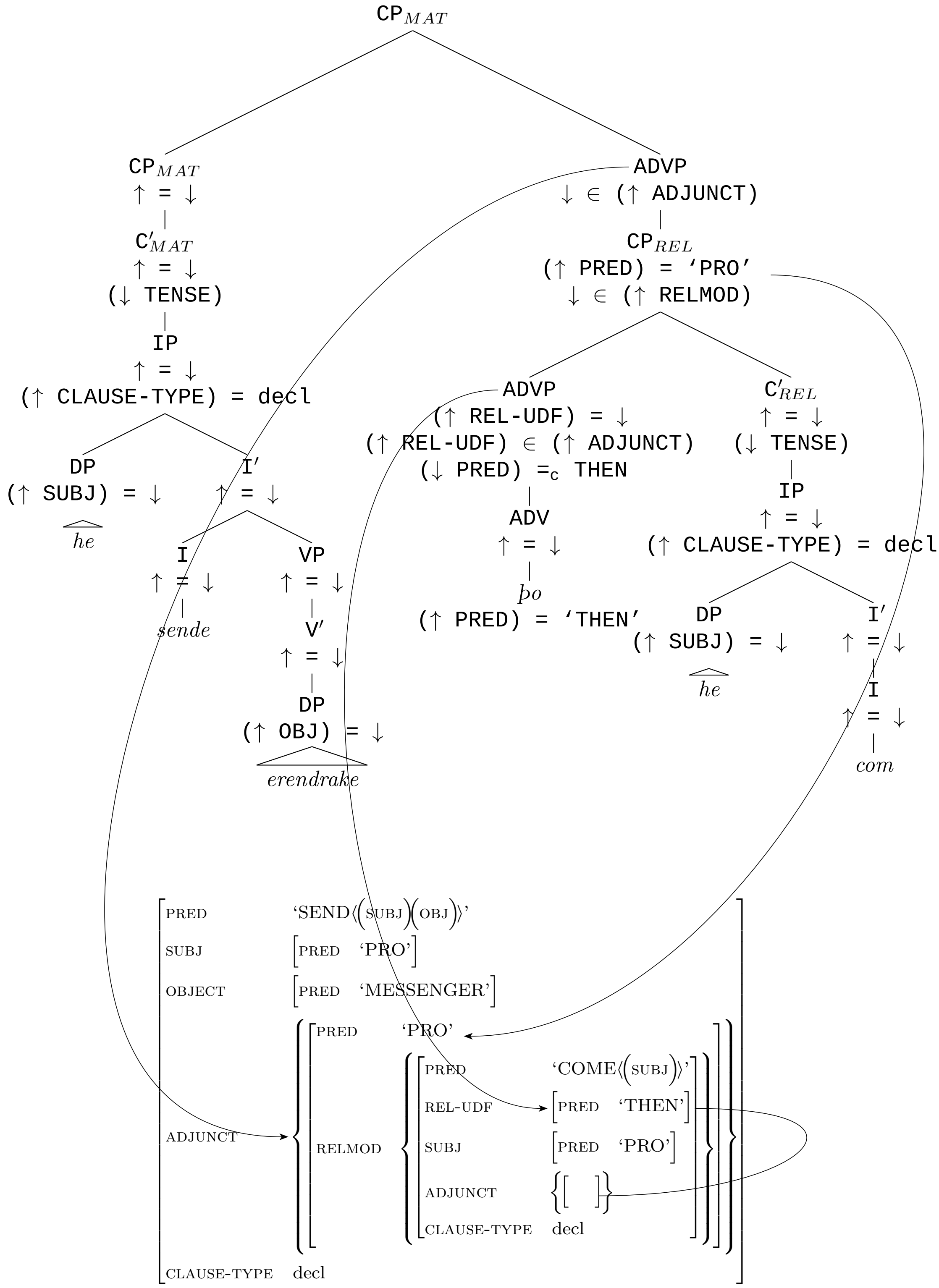
One could omit the association between relative topic and its head from the syntactic component, or implement it by means of a formal feature, say requiring a *wh*-based adverb, or by restricting the applicability of the rule directly to an instance of the lexeme *when.* Here, I endow the relative topic with a constraint checking for a *wh*-feature, `ADV-TYPE)=wh`, so that the rule could in principle be employed for all forms of relative clauses introduced by *wh*-adverbs (e.g., *come to where I live*, *the reason why you came*, *go whence you came*), but also include a more specific constraint forcing the valuation of the relative topic's `PRED`-feature as a form of *when* because this is the only item relevant for this study.

(40) $\texttt{CP}_{REL} \rightarrow$

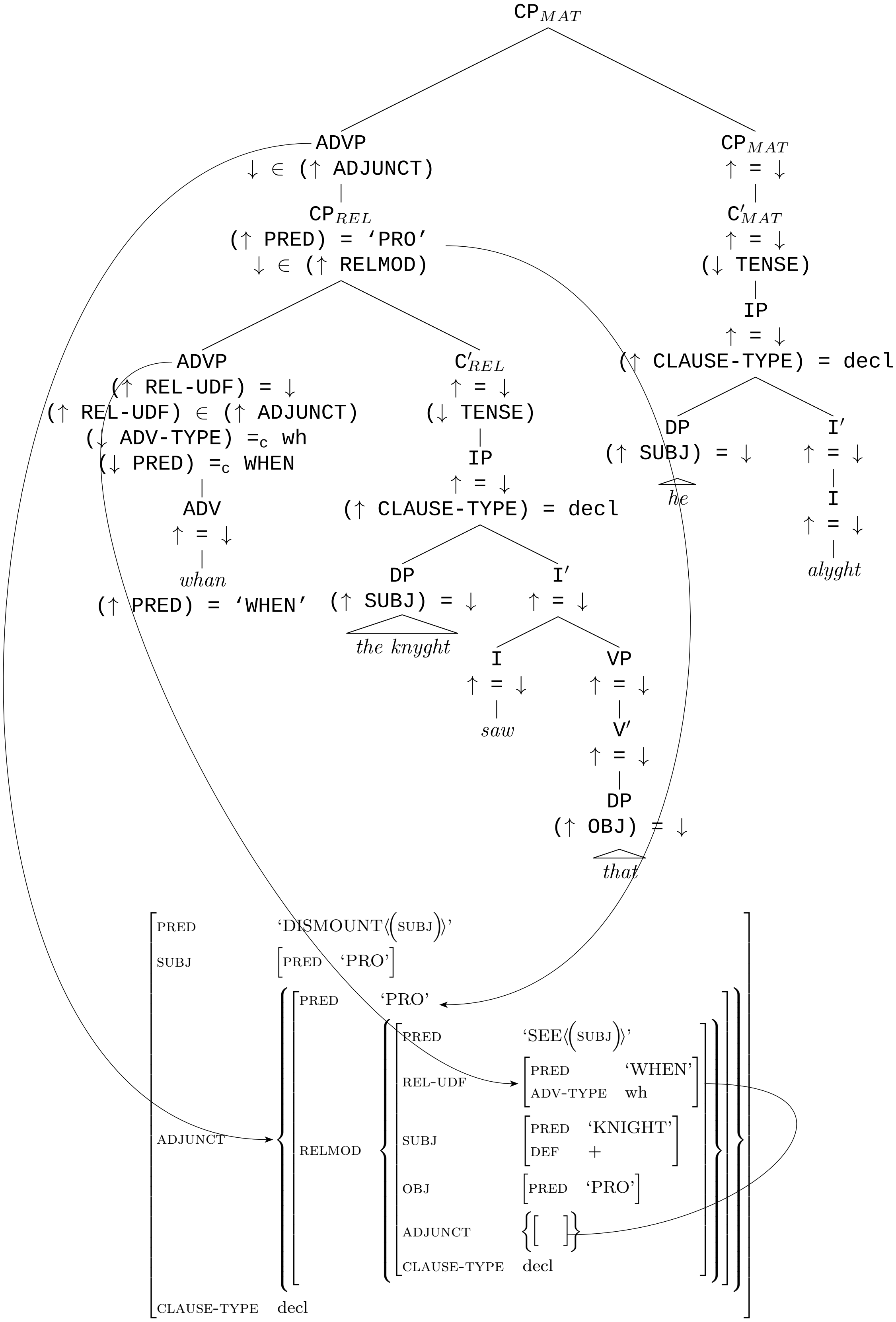
| | |
|---|---|
| ADVP | $\texttt{C'}_{REL}$ |
| $(\uparrow$ REL-UDF$) = \downarrow$ | $\uparrow = \downarrow$ |
| $(\uparrow$ REL-UDF$) \in (\uparrow \{$COMP\|XCOMP$\}^*$ ADJUNCT$)$ | $(\downarrow$ TENSE$)$ |
| $(\downarrow$ ADV-TYPE$) =_c$ wh | |
| $(\downarrow$ PRED$) =_c$ WHEN | |

My model can now parse structures that involve the conservative variant of temporal subordination based on *then*, as well as sentences with the innovative variant of temporal subordination formed with *when.* I illustrate the workings of my model with two example sentences shown in (41) for *then* and (42) for *when.* I sketch rudimentary f-structures for these sentences that contain the most important features only.

(41) he sende erendrake þo he com
he sent errand-runner then he came
'He sent a messenger when he came' (CMTRINIT,35.479)



(42) Whan the knyght saw that he alyght
when the knight saw that he dismounted
'When the knight saw that, he dismounted' (CMMALORY,40.1296)



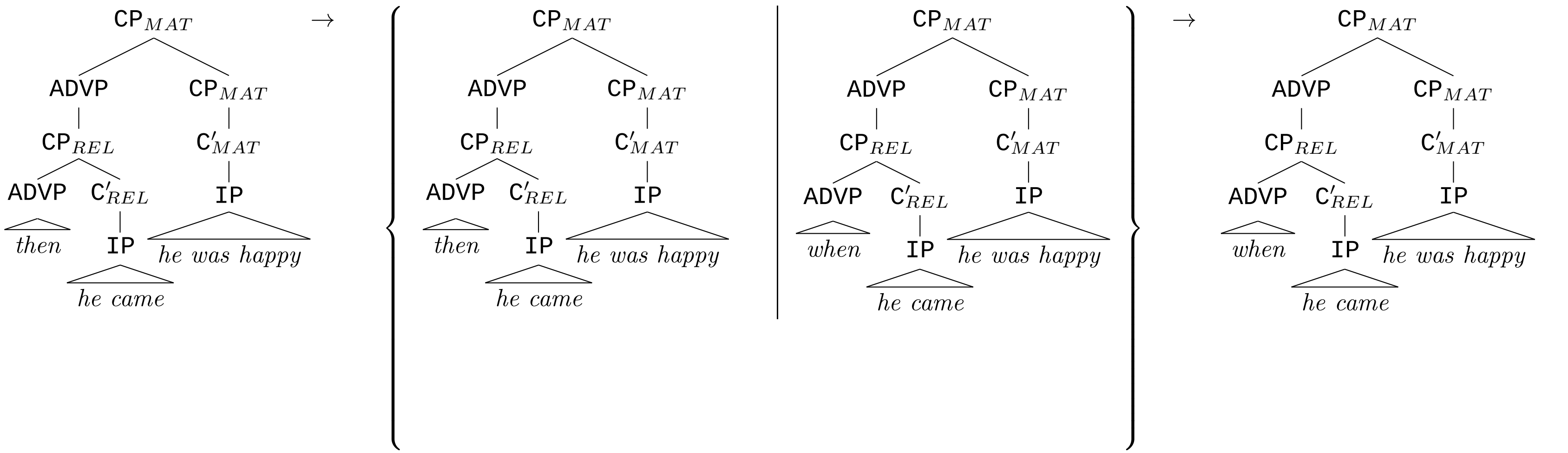
The free relatives *þo he com* ‘when he came’ and *whan the knyght saw that* ‘when the knight saw that’ are formally realized as clause-final respectively clause-initial ADVPs. They both map onto a function that is a member of the adjunct set of the matrix clause that they are attached to. This mapping is illustrated with an arrow pointing from the framing ADVP-label to the main clause’s adjunct set. The head of this adjunct member is not overtly realized, but instead gets its pronominal PRED feature directly from the phrase structure rule re-writing ADVP to $\mathrm{CP}_{REL}$. Its meaning is not given here. (It might denote some kind of underspecified relation to an event.) The introduction of the empty pronoun head is indicated with an arrow leading from the relative CP-node to the head of the adjunct. The empty pronoun head is then modified by the relative clause, RELMOD. It gets co-indexed with the relative topic, REL-UDF, in the specifier of $\mathrm{CP}_{REL}$. The specific co-indexing mechanism is not explained here. (It should lead to an expression that asserts that an event took place at a time $x$ such that he came at $x$ for the first example, or such that the knight saw that at $x$ for the second example.) In the first sentence, the dependency between the pronominal antecedent and the relative topic is established with the anaphoric item *then*, in the second sentence with the *wh*-element *when*. The mapping of these types of adverbs onto the relative topic is highlighted by an arrow from the specifier of $\mathrm{CP}_{REL}$ to the underspecified discourse function in the f-structure of the relative modifier. Finally, the underspecified discourse function of the relative clause is unified with a member of the adjunct set of the relative clause, as shown by a line ranging between these two functions.

One of the main advantages of my model lies in its parsimony. The same lexicon entries for *then* and *when* can be employed in all of their different functions. Hence, both conservative operator adverb and innovative sentence adverb *then* can also be used as relative topics to create temporal subordinate clauses. Likewise,the item *when* is represented as only one syntactic form that can be used to form direct questions as well as free relative clauses. Unnecessary doubling of lexical items is thus avoided.

The conservative rule generating temporal relative clauses with *then*, (38), enters into competition with its innovative analogue handling temporal subordinate clauses based on *when*, (40). By the second half of the fourteenth century, the latter rule has more or less completely replaced the former, at least in so far as one can judge actual usage based on written evidence. Thus, my mode also incorporates the main development of interest in this chapter, the change from subordinating *then* to *when*. The change is represented, once more, in terms of the general format for the implementation of a linguistics change with competing variants in a population of speakers from the introductory chapter.

**Change in Subordination from *then* to *when***

| Conservative period: Speakers represent only the conservative variant $\alpha$ | Transitional period: Speakers represent both the conservative variant $\alpha$ and the innovative variant $\beta$ | Innovative period: Speakers represent only the innovative variant $\beta$ |
|---|---|---|
| $\alpha \rightarrow$ | $\{ \alpha \mid \beta \} \rightarrow$ | $\beta$ |
| Conservative period: Speakers represent only a rule for the introduction of *then* into temporal relative clauses, before c. 1125 | Transitional period: Speakers represent rules for both the introduction of *then* and *when* into temporal relative clauses, c. 1125 - 1375 | Innovative period: Speakers represent only a rule for the introduction of *when* into temporal relative clauses, after c. 1375 |



#### 4.4.2.3 Related Structures

I will now discuss three additional syntactic pattens. These are (i) correlative constructions, (ii) headed temporal relatives and (iii) temporal clauses with overt complementizers.

(i) **Correlative constructions**. So-called Early English correlative constructions involve an initial temporal subordinate clause locating an event in time and discourse and a subsequent, resumptive adverb *then* establishing the reference time relative to the event expressed in the main clause (e.g., Kemenade and Los 2006*a*, Links and Kemande 2013; for a short historical summary, see Fischer et al. 2000: 88-9). As it turns out, my model can already parse such structures without any further modifications. The temporal dependent clause lacks a feature that would permit its placement in the specifier of the matrix clause `CP` and must therefore occur as an adjunct to it by rule (28). Provided that the resumptive adverb *then* includes a presentational link feature triggering verb-second order, it will be placed in the specifier of the main clause `CP`. Otherwise it will be adjoined to the matrix `CP` as well and the finite verb will remain inside the `IP`. In either case, the f-structures of the resumptive adverb and the temporal free relative will occur as members of the same main clause adjunct set. They can then simply be co-indexed with each other according to a binding principle not explicated here. Typical examples are shown in (43) with a sketch of the corresponding f-structure in (44), where the initial *then*-clause and its resumptive adverb bear the same index.

(43) a. [$_{\text{CP}}$ [$_{\text{ADVP}}$ **þa** he lai an slep in scip, ] [$_{\text{CP}}$ [$_{\text{ADVP}}$ **þa** ] [$_{\text{C}'}$ þestrede [$_{\text{IP}}$ þe dæi ouer al landes ]]]
then he lay in sleep in ship then darkened the day over all lands

‘When he lay asleep in the ship, then the day became dark in all the lands’
(CMPETERB2,54.380)

b. [$_{\text{CP}}$ [$_{\text{ADVP}}$ **wane** þi lust is ago, ]
when your lust is gone

[$_{\text{CP}}$ [$_{\text{ADVP}}$ **þonne** [$_{\text{C}'}$ is [$_{\text{IP}}$ þi song ago also ]]]
then is your song gone also

‘When your desire has passed, your song has passed also’ (OwlNight,46.508.296)

(44)
$$\left[\begin{array}{ll} \ldots & \\ \textsc{adjunct} & \left\{ \begin{array}{l} \left[\begin{array}{ll} \textsc{pred} & \text{‘PRO’} \\ \textsc{relmod} & \left\{\left[\ \ldots\ \right]\right\} \end{array}\right]_i \\ \left[\begin{array}{ll} \textsc{pred} & \text{‘THEN’} \\ \textsc{pres-link} & + \end{array}\right]_i \end{array}\right\} \end{array}\right]$$

(ii) **Headed temporal relatives**. There are quite a lot of Old and Middle English sentences that show a temporal subordinate clauses immediately following a phrase that could be said to denote a time point. I believe that such structures can be treated as the headed analogue to headless, free relatives. In other words, an overt constituent could adopt the function of the empty `PRO`-head as the antecedent of a temporal relative modifier. First, temporal `DP`s appear to be able to be modified by temporal clauses as shown in (45). Here and in the next examples, I indicate with labeled bracketing the hypothesized parse with headed relative clauses.

(45) a. ... on [$_{\text{DP}}$ [$_{\text{DP}}$ geres utgange, ] [$_{\text{CP}_{REL}}$ **þonne** þu gegaderast þine wæstmas ]]
on year’s out-going then you gather your fruits

‘... at the end of the year when you bring in your harvest’ (cootest,Exod:23.16.3324)

b. ... on [$_{\text{DP}}$ [$_{\text{DP}}$ ungewederan ] [$_{\text{CP}_{REL}}$ **þa** man oððe tilian sceolde. oððe eft tilða gegaderian ]]
in un-weather when one either till should or again tilth gather

‘... [because of] the storms when one should till or afterwards gather the fruit of the tilth’
(Chronicle4:ChronE_[Plummer]:1097.29.3292)

c. ... in [$_{\text{DP}}$ [$_{\text{DP}}$ a nyȝt ] [$_{\text{CP}_{REL}}$ **whanne** he hadde all i-seide ]]
in a night when he had all said

‘... during a night when he had said everything’ (CMPOLYCH,VI,429.3131)

Secondly, there are substantial numbers of *then then* and *then when* collocates in Old and Middle English. These structures could be said to involve the adverb *then* as the overt counterpart to the zero `PRO`-head for modification by a temporal relative clause. Examples are given in (46).

(46) a. [$_{\text{ADVP}}$ [$_{\text{ADVP}}$ þa ] [$_{\text{CP}_{REL}}$ **þa** me hine to beheafdunge lædde. ... ]]
then then one him to beheading led

'When he was led to his beheading ...' (cojames,LS_11_[James]:110.102)

b. [$_{\text{ADVP}}$ [$_{\text{ADVP}}$ þanne ] [$_{\text{CP}_{REL}}$ **huanne** we ziggeþ 'vader oure' ... ]]
then when we say father our

'When we say the Lord's Prayer ...' (CMAYENBI,101.1985)

Following this line of reasoning, one would be led to the assumption that other time point adverbs can be modified by *then-* or *when*-clauses as well. Potential examples of this sort are presented in in (47).

(47) a. [$_{\text{ADVP}}$ [$_{\text{ADVP}}$ Nu ] [$_{\text{CP}_{REL}}$ **þenne** hit wat Gorlois hu hit iuaren is ... ]]
now then it knows Gorlois how it fared is

'Now that Gorlois knows how it has fared ...' (LAYAMON,496.260)

b. And [$_{\text{ADVP}}$ [$_{\text{ADVP}}$ after, ] [$_{\text{CP}_{REL}}$ **whan** he was Crouned, ...]]
and after when he was crowned

'And afterwards, when he was crowned, ...' (CMBRUT3,12.323)

For all of these structures, it is not clear whether a temporal phrase, say as *the day*, *then* or *now*, preceding a temporal subordinate clause, for instance *when I go out*, really functions as its antecedent, as sketched by the rudimentary f-structure in (48a), or if they are two independent structures, a nominal/adverbial adjunct and a temporal free relative respectively, modifying the same mother node, as outlined by the skeleton f-structure in (48b).

(48) a.
$$\left[\begin{array}{ll} \ldots & \\ \text{ADJUNCT} & \left\{\left[\begin{array}{ll}\text{PRED} & \text{`DAY/THEN/NOW'}\\ \text{RELMOD} & \left\{\left[\ \ldots\ \right]\right\}\end{array}\right]\right\}\end{array}\right]$$

b.
$$\left[\begin{array}{ll} \ldots & \\ \text{ADJUNCT} & \left\{\begin{array}{l}\left[\begin{array}{ll}\text{PRED} & \text{`DAY/THEN/NOW'}\end{array}\right]\\ \left[\begin{array}{ll}\text{PRED} & \text{`PRO'}\\ \text{RELMOD} & \left\{\left[\ \ldots\ \right]\right\}\end{array}\right]\end{array}\right\}\end{array}\right]$$

It might be possible to decide between the two options on a case by case basis. The temporal clause could be analyzed as an ordinary or a free relative clause if its relation to the preceding phrase feels more restrictive or more appositive respectively. Furthermore, one could test if relativization can be achieved without any relativzers at all (*the day you saw her you fell in love*, vs. **later you saw her you fell in love*). In general, however, the structures are probably genuinely ambiguous between headed and free relative analyses. This ambiguity may also be reflected in variable annotations in the parsed corpora used here (e.g., *[a tyme [whan kynge Arthure was at London]]* parsed as a headed relative clause, CMMALORY,45.1483 vs. *[þe fyrst tyme] [when a man es turned to God]* parsed as two independent adjuncts, CMROLLEP,78.245).

I will therefore add rules to my model that can generate headed temporal relative clauses. Specifically, I will include alternatives to re-writing a syntactic label to an empty pronoun antecedent and a relative clause (see e.g., (35)), and allow the syntactic label to derive instead an overt temporal determiner phrase, say *the day*, (49a), or an overt temporal adverb phrase, say *now*, (49b) followed by a relative clause, say, *... when something happens*. Hence, the antecedent of the relative clause is not headed by a zero pronoun, but instead gets its `PRED`-feature directly from the overt `DP` or `ADVP`. I will not include conditions on the establishment of a permissible binding relationship between the antecedent and the relative topic. This would require a formalization of the denotation of temporal relative operators and their selectional requirements for an appropriate temporal argument (*[the time]$_i$ when$_i$ I slept _*, **[the house]$_i$ when$_i$ I slept _*; *[now]$_i$ when$_i$ people go out _*, **[badly]$_i$ when$_i$ people go out _* etc.). Instead, I will leave the syntactic rules completely underspecified in this respect.

(49)
$$\begin{array}{llll} \text{a.} & \mathtt{DP} \rightarrow & \mathtt{DP} & \mathtt{CP}_{REL} \\ & & \uparrow = \downarrow & \downarrow \in (\uparrow \mathtt{RELMOD}) \\ \text{b.} & \mathtt{ADVP} \rightarrow & \mathtt{ADVP} & \mathtt{CP}_{REL} \\ & & \uparrow = \downarrow & \downarrow \in (\uparrow \mathtt{RELMOD}) \end{array}$$

These rules will now permit my model to provide a structural analysis for sentences involving modification of an overt nominal or adverbial antecedent with a temporal relative clause.

(iii) **Overt complementizers**. Finally, there are Old and Middle English temporal subordinate clauses with overt C-heads. In Modern English, the co-occurrence of a *wh*-based relative topic and a `C`-head is generally precluded, but in earlier stages of the language, this restriction, commonly called the doubly filled COMP filter, applied much less rigorously. In particular, relativizers, such as Old English *þe* or Middle English *þat*, could appear alongside subordinating *then* or *when* in temporal free relative constructions. This is illustrated in (50) below.

(50) a. and [$_{\text{CP}}$ þa [$_{\text{C}'}$ **þe** [$_{\text{IP}}$ he wæs twelf wintra ]]] he gewilnode to westene
and then that he was twelve winters he wanted to wasteland

'And when he was twelve years old he wanted [to go] to a desert'
(coaelive,ÆLS_[Martin]:23.5986)

b. [$_{\text{CP}}$ ðonne [$_{\text{C}'}$ **þæt** [$_{\text{IP}}$ man þone wyrttruman cnucige & to ðam sare gelecge ]]]
then that one the herb-roots pounds and to the sore lays

eal þæt sar he gehælð
all the sore it heals

'When one pounds the roots of this herb and puts it on a sore spot, it will heal all the soreness'
(coherbar,Lch_I_[Herb]:51.2.1017)

c. And [$_{\text{CP}}$ whan [$_{\text{C}'}$ **that** [$_{\text{IP}}$ they knewe that they were naked, ]]]
and when that they knew that they were naked

they sowed of fige leves a maner of breches
they sewed of fig leaves a manner of breeks

'And when they realized that they were naked, they made some kind of pants from fig leaves'
(CMCTPARS,297.C1.369)

In order to parse sentences such as the above, I will develop a rule that allows the C-position of a $\texttt{CP}_{REL}$ to be projected. The lexicon entry for the relativizer *that* is shown in (51). The featural contribution of this item can be left entirely unspecified for my model to produce adequate parses and I will therefore leave the item as a vacuous formative. It is, however, possible that the relativizer reinforces certain constraints imposed on the relative clause, such as the requirement to be finite or to occur with a declarative clause type feature and, in this respect, aids in sentence processing.

(51) *that* $\texttt{C}_{REL}$

The required syntactic rule can then simply re-write $\texttt{C}'_{REL}$ to $\texttt{C}_{REL}$ and its complement `IP`. As before, the rule will add a declarative clause type feature, (52).

(52)
$$\texttt{C}'_{REL} \rightarrow \begin{array}{cc} \texttt{C}_{REL} & \texttt{IP} \\ \uparrow = \downarrow & \uparrow = \downarrow \\ & (\uparrow \texttt{CLAUSE-TYPE}) = \texttt{decl} \end{array}$$

The inclusion of this rule will now allow my model to parse temporal relative clauses with overt C-heads. This concludes my formalization of the Old and Middle English items *then* and *when*.[5]

### 4.4.3 Predictions

The general hypothesis to be tested in this chapter is that the development from verb-second order after operator adverb *then* to 'subject - verb' order after sentential modifier *then*, as formalized in the first section of the preceding part, constitutes an antecedent cause for the change from subordinating *then* to subordinating *when*, as formalized in the second section of the preceding part. In order to demonstrate that this explanation has some merit, one must examine consequences that should be true if the hypothesis is true and that should be false if the hypothesis is false. Unfortunately, it is usually very difficult to conceive of predictions that follow specifically from hypothesized causes of a linguistic change. The present case is no exception. In fact, I could think of only two predictions that could be informative about the causal relation between the syntactic word order change and the innovation in subordinator form.

[5]The grammar fragment has been implemented as an LFG toy grammar that can be run on the Xerox Linguistic Environment (XLE) grammar engineering platform. It is stored on the Dissertation DVD under `Chapter 4 - WhenThen/02 Toy Grammar/Ch4WhenThen.lfg`.

The first prediction is the following. If it is indeed true that the loss of verb-second order after *then* causes the rise in subordinating *when*, then the former change should occur earlier than the latter. This follows simply from the assumption that a cause has to precede its effect in time. Moreover, the starting point of the syntactic change should not lag far behind the onset of the change in morphological realization of the subordinator. The reason for this is that causes and effects usually occur in close temporal proximity to each other, at least where the cause is a sufficient one. The hypothesis test will only be meaningful if the corpus texts adequately reflect the relative order of linguistic changes. I will assume that the texts in my sample meet this requirement.
The prediction can then be investigated as follows. The two supposedly associated linguistic changes have to be measured independently and then be put into a temporal relation to one another. The decline in verb-second order after *then* can be measured as the proportion of main clauses with the finite verb immediately following initial adverbial *then* out of all main clauses featuring initial adverbial *then.* The rise in *when* can be measured as before. Logistic regression models predicting the variable patterns from time should then show an earlier point of origin for the former than for the latter change. Further, one should find that the syntactic change does not progress for very long before the inception point of the change in subordinator form.

(53) **Hypothesis 1:** *Time Course of Decline in Verb-Second after* then *and Rise in* when
If one measures the decline in verb-second order after *then* and the rise in subordinating *when*, one should find that the former change commences earlier than, but does not become implemented excessively before, the onset of the latter change.

The second prediction reads as follows. If the decline in subordinating *then* is due to the loss of syntactic word order clues for its interpretation, then the declining effect should be reduced if the subordinate reading is signaled through other means. In other words, the incidence of temporal *then*-clauses should survive more robustly where they involve alternative subordinating strategies. These alternative strategies were listed in the third section of the preceding part: (i) correlative *þa ... þa* constructions, (ii) a salient overt antecedent for the temporal relative clause, such as *then*, hence *then then* or *then when* patterns, and (ii) the presence of an overt complementizer in a free relative clause.
The hypothesis can be tested by coding every example of subordinating *then* and *when* for the presence or absence of an alternative subordinating strategy. A logistic regression model that predicts the realization of the subordinator from time and the strategy variable should then show an overall lower probability of *when* if an independent signal of subordination is present. This could be interpreted to mean that the overt indication of subordination protects conservative *then* from its innovative competitor.

(54) **Hypothesis 2:** *Effect of Alternative Subordinating Strategies*
If the temporal subordination data is subdivided into those cases that involve an alternative indication of embedding and those cases that do not, one should discover an overall lower probability of *when* in the former than the latter context.

I would like to stress again that the results of the tests of these hypotheses will not be able to establish beyond reasonable doubt that the investigated word order changes must exert a causal effect on the change in formal expression of temporal subordination. First, other factors, such as French calquing, analogy, etc. could be more decisive than the syntactic developments. A detailed model considering such influences might therefore show that the explanatory value of the word order variable could disappear or become much less important than when its effect is studied in isolation. Secondly, the results of the above predictions themselves will be relatively indeterminate. A positive result of the time course argument would merely demonstrate the possibility of a causal connection between syntax and subordinator form, but not the plausibility of such a link. Likewise, the argument from the second hypothesis remains relatively weak as I am not able to explicitly spell out the actual mechanisms from which the intuitively plausible effect is supposed to follow. Explanations along evolutionary or functional lines given earlier may be possible, but these explanations are themselves fairly speculative. Other researchers may be able to interpret the outcome of the study - in whichever direction it might go - in ways that might not necessarily inform the question of the causal role of word order signals for the change in temporal subordination. Nevertheless, the investigation of these hypotheses is important because it will provide at least some evidence for or against the role of word order cues as a reason for the formal change in clause type markers. Limited evidence for a claim is obviously preferable to mere speculation without any empirical corroboration whatsoever.

### 4.4.4 Summary

In this part, I offered speculations on the cause of the Middle English replacement of *then* with *when* in their function as temporal subordinators. I identified three factors to which it seemed plausible to me that one could attribute a causal role for this development: French borrowing, analogical extensions of relative *wh*-items, and the loss of verb-second word order signaling an adverbial reading of *then*. Focusing on the third influencing factor, I then developed an LFG toy grammar that formalizes the decline in verb-second patterns after adverbial *then*, the rise in *when*-based subordination, as well as some related structures. The change from *then* to *when* was modeled as a small alternation of a constraint annotated on a phrase structure rule responsible for generating temporal adverbial relativization. Finally, I derived two quantitatively testable predictions from the contention that the decline in verb second order after *then* is causally related to the increase in subordinating *when*.

## 4.5 Hypothesis Testing

I will now put the hypotheses derived in the previous section to the test. For each of the two hypotheses, I will first describe the data collection process and then evaluate the results.

### 4.5.1 Hypothesis 1: Time Course of Decline in Verb-Second after *then* and Rise in *when*

According to my first hypothesis, verb second order should decline earlier than, but should not have progressed very far by the time of, the rise in *wh*-based subordination. This assumption can be investigated straightforwardly by measuring the syntactic patterns after adverbial *then* and compare the finding to the progression of the rise in subordinating *when*.

#### 4.5.1.1 Data Collection

In principle, it is an easy task to measure the decline in verb-second patterns after the adverb *then*. One would simply have to compare 'then - verb - subject' to 'then - subject -verb' structures.

In practice, however, an adequate measurement of the decline in verb-second order after *then* is complicated by a number of confounding factors. (i) The conservative structure becomes fossilized to a certain degree. Hence, the finite verb can still follow *then* even today in archaic, lyrical or dramatic settings. This seems to be the case, in particular, with the unaccusative verb *come*. One might thus expect to find significant genre, verb type or lexical effects. (ii) The probability of verb second vs. subject-verb patterns after *then* may not just be a cue for its interpretation, but may itself be influenced by syntactic word orders. For example, if *then* appears after a temporal clause in a correlative construction, it may receive an independent signal for its function as a main clause adverb, thus lowering the probability that it will be followed by verb-second order. Syntactic context could thus have a significant impact on word order after *then*. (iii) Most importantly, there is a general increase in high verb placement under C during the Middle English period. This is the only factor that I will control for in this study to some degree. I will first describe the phenomenon and why it poses an obstacle for the intended measurement. Subsequently, I shall explain how I attempted to circumvent the problem.

Old English displayed V-to-C structures, hence potential verb placement before subject pronouns, only in a restricted number of contexts, namely in direct questions, after operator adverbs, under negative polarity and with non-indicative mood (see section 3.3.1.4 of the previous chapter). Middle English, however, often shows high verb placement, hence possible inversion with pronominal subjects, after any kind of initial constituent in declarative matrix clauses. Examples with initial object topics are presented in (55). They are all taken from the East Midlands dialect, which is the most direct Middle English ancestor of the later, London-based standard variety of English.

(55) a. Chaucer's *Parson's Tale*, c. 1390

And [$_{\text{CP}}$ moore shame [$_{\text{C}'}$ *do* [$_{\text{IP}}$ **they** to Crist, than dide they that hym crucifiede, ]]]
and more shame do they to Christ than did they that him crucified

'And they do more shame to Christ than do those who crucified him' (CMCTPARS,309.C1.891)

b. *The Book of Vices and Virtues*, c. 1400

and [$_{\text{CP}}$ many wordes [$_{\text{C}'}$ *fynde* [$_{\text{IP}}$ **we** þat scheweþ vs wel what is of him ]]]
and many words find we that show us well what is of him

'And we find many words that show us well what it is about him' (CMVICES4,102.112)

c. *The Book of Margery Kempe* c. 1435

“Nay,” he seyd, “[$_{\text{CP}}$ þat [$_{\text{C}'}$ *wyl* [$_{\text{IP}}$ **I** not grawnt ʒow ]]]”
no he said that will I not grant you

“‘No,” he said, ”That, I will not grant you.” (CMKEMPE,24.504)

A change therefore seems to have been actuated during Middle English that put the language on a trajectory towards a generalized verb-second constraint, of the type found in some Modern West-Germanic languages, but that was never carried out completely and was eventually reverted.

This peculiar stage in the evolution of English verb positions has received a fair amount of attention in the specialized literature. For example, Haeberli (2007)’s text sample shows inversion of finite verb and pronominal subject after fronted elements at substantial rates of 18.5% in Helsinki period M3 or 38.1% in period Helsinki M3/4. He also shows that the difference between transitive verbs and unaccusatives / *be* may have a substantial effect on verb placement and finds that there is tremendous variation among individual authors, with some avoiding inversion completely and others inverting subject and finite verb over 50% of the time. It is also known that there is an important dialect split such that more Northern Middle English texts have a greater probability of displaying high verb placement after initial constituents than more Southern Middle English texts (Kroch and Taylor 1997). For example the Northern *Prose Rule of St. Benet* shows almost categorical inversion after any fronted constituents with both full and pronominal subjects (ibid.: 313). The dialect effect was confirmed by Trips (2002: ch. 6.4.), who shows competition between verb-second and non-verb-second grammar options in Middle English texts, providing, among others, a detailed investigation of Richard Rolle’s writings. Studies on the historical development of inversion after the Middle English period are summarized in Fischer et al. (2000: 132–4). Referencing a study by Nevalainen (1997), the authors suggest that inversion may have declined from a high of c. 37% to under 10% between 1420 and 1710. Hence inversion after initial constituents does not reach low rates comparable to the present-day stage of the language until well into the 18$^{\text{th}}$ century.

The rise and fall of high verb placement during the time of late medieval and Renaissance English causes ambiguity for the proper analysis of Middle English instances of verb-second order after adverbial *then* as in (56). (i) Such sentences may be indicative of the targeted structure with an Old English-type operator adverb. In this case, a presentational link feature on *then* would be parsed with a rule constraining the clause to be a non-canonical declarative and hence occur with the verb under C. (ii) However, these structures could also be parsed as confounding instances of generalized inversion after any initial constituent. In that case, the featural make-up of *then* would be irrelevant for the syntactic position of the verb.

(56) [$_{\text{CP}}$ **þan** [$_{\text{C}'}$ *enters* [$_{\text{CP}}$ þou into þe toþer degre of lufe ]]]
then enter you into the other degree of love

‘Then, you will enter into the other degree of love’ (CMROLLEP,64.71)

(i) *then* as operator adverb attracts the verb?
or: (ii) the verb is free to be placed high after any initial constituent?

How could one overcome the problem of the indeterminate source for high verb placement after *then* in Middle English? My attempted solution is to modify the proportion of observed verb-second structures after *then*. Specifically, I will reduce the number of observed ‘then - verb’ cases according to an independent measurement of general ‘X - verb’ structures for every text. I will refer to a text’s correcting factor for the adjustment of the incidence of verb-second orders after *then* as its ‘V-to-C index.’

The index for generalized V-to-C structures was established as follows. I decided to trace the development of inversion across the entire history of English. Therefore, this study not only considered the Old and Middle English texts described earlier, but also material from the Penn-Parsed corpora of early Modern English (Kroch et al. 2004) (release with Helsinki, Penn1 and Penn2 subsections, 449 files) and Modern British English (Kroch et al. 2010) (release 1 with 101 text files). The dates of composition of early Modern and Modern English texts are quite certain. They can be found on the websites of the respective corpora and I will therefore not list them here again. I collected all sentences with an initial, topicalized, full DP object. For the purpose of this study, such objects were defined as the set of the nodes `NP|NP-ACC|NP-DAT` with or without a hashtag indicating displacement dominating at least one item from the set `N^*| NR^*| *Q^*| D^*| N| NR| *Q| D| NUM*` for Old English, and as the set of nodes `NP-OB1|NP-OB2|NP-#` dominating at least one member of the set `N*| *+N*| *D*| *Q*| *ONE*| *NUM*` for the other language stages. I only considered pronominal subjects for this study. Pronoun subjects were defined as in section 3.4.2, i.e., as one-word subject nodes dominating only a personal pronoun, the indefinite pronoun *man* or a demonstrative. Finite verbs were restricted to positive, potentially indicative forms for Old English and positive verbs for the other periods. In this way, I ruled out that verb placement before subject pronouns was licensed by negation or non-indicative mood. Then, I collected all examples of two word order patterns. The object topic could either be immediately followed by the finite verb, 'object - verb - subject,' indicating a V-to-C-structure, or it could be followed at any distance by the subject and subsequently the verb, 'object - subject - verb,' indicating placement of the verb inside the IP. The object topic did not have to be the first element in the clause, and thus could be preceded by other elements, such as vocatives, conjunctions, subordinate clauses etc.

Examples from the four periods of the history of English are presented below.

(57) **Old English**

a. *Object - Verb ... Pronominal subject* (19 examples)

& [$_{CP}$ **mycel ærende** [$_{C'}$ *brohte* [$_{IP}$ he ]]]
and great errand brought he

'And he brought a great message' (coblick,HomU_18_[BlHom_1]:9.101.100)

b. *Object ... Pronominal subject ... Verb* (1340 examples)

[$_{CP}$ **þone wineard** [$_{CP}$ [$_{IP}$ he *behegode* ]]
the vineyard he behedged

'He built a hedge around the vineyard' (coaelhom,ÆHom_3:81.458)

(58) **Middle English**

a. *Object - Verb ... Pronominal subject* (408 examples)

And þus [$_{CP}$ **þe comun vndyrstondyng** [$_{C'}$ *schulden* [$_{IP}$ we algatis holde ]]]
and thus the common understanding should we all-ways hold

'And thus we should hild the common understanding in every respect' (CMWYCSER,348.2183)

b. *Object ... Pronominal subject ... Verb* (1280 examples)

[$_{CP}$ **Syche dowtis** [$_{CP}$ we *schulden* sende to þe scole of Oxenforde ]]
such doubts we should sent to the school of Oxford

'We should send such doubts to the school at Oxford' (CMWYCSER,370.2585)

(59) **Early Modern English**

a. *Object - Verb ... Pronominal subject* (190 examples)
And [$_{CP}$ **these rules** [$_{C'}$ *have* [$_{IP}$ I here set down rather to informe the less skilful teacher ]]]
(HOOLE-E3-P1,18.68)

b. *Object ... Pronominal subject ... Verb* (793 examples)
[$_{CP}$ **a questyon** [$_{IP}$ I *have* here to move unto yow, wheryn I shall dyssyer yow to be playne ]]
(MOWNTAYNE-E1-P2,197.49)

(60) **Late Modern (British) English**

a. *Object - Verb ... Pronominal subject* (30 examples)
Damn the word [$_{CP}$ **more** [$_{C'}$ *will* [$_{IP}$ I tell you ]]]
(COLMAN-1805,28.315)

b. *Object ... Pronominal subject ... Verb* (220 examples)
[$_{CP}$ **One Thing more** [$_{IP}$ I *will* only premise ]]
(DODDRIDGE-1747,29.234)

The purpose of this study was to find out how frequently a finite verb would occur under C during Middle English. I therefore did not consider in detail any factors that would be relevant for the realization of this syntactic pattern in later periods. In particular, the Early Modern English rise of *do*-support should lead to a situation in which only auxiliaries can invert with the subject, but I did not make a distinction between lexical and auxiliary verbs here.

I conducted a total of 2 (word order patterns) * 4 (period) = 8 search queries for the determination of the V-to-C index.[6]

[6]The CorpusSearch query and output files as well as an excel summary sheet for the calculation of the 'V-to-C index are saved on the Dissertation DVD under `Chapter 4 - WhenThen/03 Hypothesis Testing/H1/V-to-C-Index`.

Next, I proceeded to measure the decline in verb-second order after adverbial *then*. For this purpose, I counted all main clauses in which the finite verb immediately follows an instance of the lexeme *then*, as in (61), and in which the verb follows an overt subject in this context instead, as in (62). While potentially relevant, I did not take any other factors into account. For example, I considered all verb types jointly, such as unaccusatives, illustrated in the a.-examples, transitives, as in the b.-examples, modal auxiliaries, the c.-examples, or copulas, as in the d.-examples. Likewise, the measurement did not distinguish between different subjects types, like full nominal subjects, shown in the a.-examples, plural, b.-examples, or singular pronouns, c.-examples, or proper names, d.-examples. I also assume that the change should occur independently of genre and hence grouped together poetry, illustrated in (61c) and (61d), with prose texts, illustrated by the other examples.

(61) *Conservative variant:* then *with verb-second order:*

a. [$_{\text{CP}}$ **þen** [$_{\text{C'}}$ *come* [$_{\text{IP}}$ our lady to hur ]]]
then came our lady to her
'Then, our lady came to her' (CMMIRK,110.3014)

b. [$_{\text{CP}}$ **Đa** [$_{\text{C'}}$ *cusen* [$_{\text{CP}}$ hi an clerc ]]]
then chose they a clerk
'Then, they chose a clerk' (CMPETERB1,43.55)

c. [$_{\text{CP}}$ **þen** [$_{\text{C'}}$ *maiȝt* [$_{\text{IP}}$ þou synge of loue lele. ]]]
then might you sing of love loyal
'Then, you may sing of faithful love' (HowHearMass,134.208.69.[Stanza_18])

d. [$_{\text{CP}}$ **þo** [$_{\text{C'}}$ *was* [$_{\text{IP}}$ hauelok in ful strong pine. ]]]
then was Havelok in full strong pain
'Then, Havelok was in very strong pain' (Havelok,17.540.246)

(62) *Innovative variant:* then *followed by 'subject- verb' order:*

a. [$_{\text{CP}}$ **Then**, [$_{\text{CP}}$ on þe morow, [$_{\text{CP}}$ [$_{\text{IP}}$ mongkes *come* to hym ]]]]
then on the morrow monks came to him
'Then, the next morning, some monks came to him' (CMMIRK,100.2723)

b. and [$_{\text{CP}}$ **þan** [$_{\text{CP}}$ [$_{\text{IP}}$ þei *chose* Jon þe XXII ]]]
and then they chose John the 22
'And then, they chose [Pope] John XXII' (CMCAPCHR,141.3280)

c. [$_{\text{CP}}$ **þan** [$_{\text{CP}}$ [$_{\text{IP}}$ þou *maiste* seye, "Lorde, I haue on þi leueree" ]]]
then you may say Lord I have in your delivery
'Then, you may answer, "Lord, I have [done so] by your sustenance"' (CMROYAL,19.175)

d. [$_{\text{CP}}$ **þo** [$_{\text{CP}}$ [$_{\text{IP}}$ Stevene *was* pope sevene ȝere. ]]]
then Steven was pope seven years
'Then, Steven was pope for seven years' (CMPOLYCH,VI,429.3135)

I required *then* to be immediately adjacent to the subsequent sentence element, the finite verb for the conservative variant, the subject for the innovative variant. However, I added several syntactic categories to the ignore-list so that these categories were ignored for the calculation of adjacency conditions. To be precise, the ignore-list included the following list of categories: `*CONJ| NEG| INTJ*| CP-ADV*| PP| PP-LFD`. Thus, other adjuncts could intervene between innovative *then* and the subject, such as a `PP` in (62a). Next, I removed all instances of the expression *hwæt þa* 'see then!, Behold!' from the Old English hits because it seems to me that *hwæt* forms a constituent with *then* in these cases. Furthermore, Old English *then* had to be the first element of the clause, bar the categories of the ignore-list mentioned above as potential interveners. The reason was that I wanted to reduce the chances of retrieving cases in which *then* would be placed lower in clause, say adjoined to `VP`, co-occurring with a low subject. I lifted this restriction for Middle English and the subsequent periods, assuming that subjects can no longer remain in a lower subject position from that language stage on.

Again, I collected the conservative and innovative word orders after adverbial *then* from the entire history of English. Hence, I executed in total 2 (word order patterns) * 4 (period) = 8 search queries for the collection of the *then*-data.[7]

[7] The CorpusSearch query and output files as well as an excel summary sheet for the *then*-data can be found on the Dissertation DVD under `Chapter 4 - WhenThen/03 Hypothesis Testing/H1/`.

#### 4.5.1.2 Results

I will first evaluate the findings from my measurement of the development of V-to-C structures with object topics and pronominal subjects. I removed from the data all hits of Early Modern English and Modern English renditions of the bible (bibletyndold, bibletyndnew, bibleauthold, bibleauthnew, purverold, purvernew, ervold, ervnew) as well as Queen Elizabeth I's and Henry James' *Boethius* translations (boethel, boethja). The reason was that these texts seemed to me to employ a particularly archaizing style. This left a total of 4,051 examples. The results confirm that high verb placement under C evolves very much as has been reported in the professional literature. The order 'object topic - verb - subject pronoun' is virtually non-existent in Old English, but becomes more frequent from c. 1000 on, reaching a peak in the late 14$^{th}$ and early 15$^{th}$ centuries. Afterwards, inversion gradually declines until it virtually disappears from the language by about 1800. Hence, the V-to-C index shows an inverted u-shaped curve. At around 1400, there appears to be very great variation in the proportion of inversion among individual authors. This variability can in part be attributed to a dialectal split. Northern texts, such as the *Rule of St. Benet* or the Thornton manuscript version of the *Mirror of St. Edmund*, show a higher proportion of V-to-C patterns than most Southern writings. Similarly, elevated rates of inversion are also found in the *Ormulum* or *Havelok the Dane*, texts which show considerable Norse influence and were composed within the Danelaw, perhaps Lincolnshire. This might indicate that one can adequately measure the proportion of high verbs in poetic texts as well. On the other hand, relatively high values of the V-to-C index are also encountered in East Midlands texts, such as Chaucer's compositions, *The Cloud of Unknowing*, *The Book of Margery Kempe* or Capgrave's *Chronicle of England*, and some West-Midlands texts, notably Malory's *Morte d'Arthur*.

Figure 4.3 illustrates the development of inversion after object topics with pronoun subjects throughout the history of English. Every text in the data set is represented by a data point that shows its proportion of high verb placement and that is proportional to the number of examples it contains. Some text names are explicitly indicated. The graph also includes a smoothing line based on a local, non-parametric regression, called a Loess curve, and its 95%-confidence interval, which clearly reveals an inverse u-shaped trend in the data.

Next, I will discuss the development of high verb placement after adverbial *then*. I discarded the same dubious texts as for the previous investigation. This left 16,372 examples in the data set. I then traced the evolution of verb-second after *then* throughout the history of English and found that this pattern appears to have declined in a very gradual, drawn-out process lasting roughly from the 11$^{th}$ to the 18$^{th}$ century. I adjusted the proportion of this pattern for every text according to the findings from the previous study. Specifically, I removed from every text the percentage of 'then-verb-subject' sentences, rounded to the nearest natural number, that corresponds to its V-to-C index. For example, Malory's *Morte d'Artur* inverts pronominal subject and verb in 36 out of 60 cases with a clause-initial topicalized object, which translates into a V-to-C index of 60%. The text includes 485 instances of initial *then*, of which 97 are followed by verb-second patterns and 388 by 'subject - verb' order. Thus, I removed 60% of the former word order option, reducing its number from 97 to just 39. Consequently, the estimated proportion of genuine verb-second order after *then* drops from 20% (97 out of 485 cases) to 9% (39 out of 427 cases). This calculation reflects the consideration that a large number of *then*-initial verb-second clauses may not have resulted from the presence of an operator feature, but from a generalized verb-second constraint permitting inversion after any kind of fronted material. 242 documents did not have any examples of fronted objects and subject pronouns. The V-to-C index for these texts was set to 0, and their proportion of inversion after *then* thus remained uncorrected. The modification reduced the overall number of examples from 16,327 to 15,413. Two mixed-effects regression models including the time variable and a random text effect fitted first to the uncorrected and the the corrected data sets show a moderate reduction in the random text variance parameter from 2.251 to 1.875. While there remains substantial variability in the frequency of verb-second patterns after *then* among the texts, this finding nevertheless suggests that the correction of the counts according to the V-to-C index leads to a more realistic assessment of the change.

Figure 4.4 depicts the development of high verb placement after *then*. It shows the proportion of 'then - verb - subject' structures as a data point for every text. The gray points represent its uncorrected, the green points its corrected values. Smoothing Loess curves represent the overall trend in the data. Due to the adjustment of the incidence of verb-second patterns, the gray, uncorrected regression line runs substantially higher than the green, corrected curve in particular for the time between 1200 and 1600. High verbs after adverbial *then* have become very infrequent by the late 18$^{th}$ century so that the change can be regarded as complete by that time.

As a side note, the rate of high verb placement both after initial object topics and after initial *then* may have been increasing since the 1800s. This finding is not of actual interest for the purposes of this hypothesis test, and may turn out to be due a sampling accident. Nevertheless, the observation is interesting and may be investigated more closely in future research.

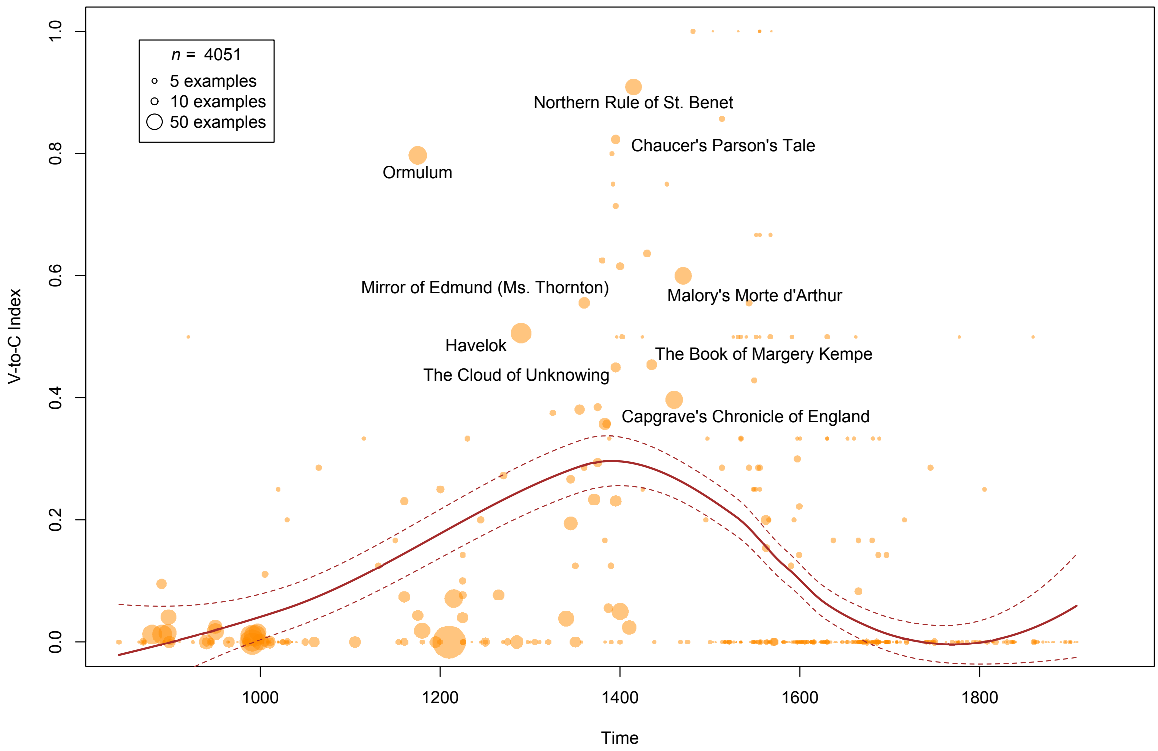



Figure 4.3: Development of the V-to-C index throughout the history of English

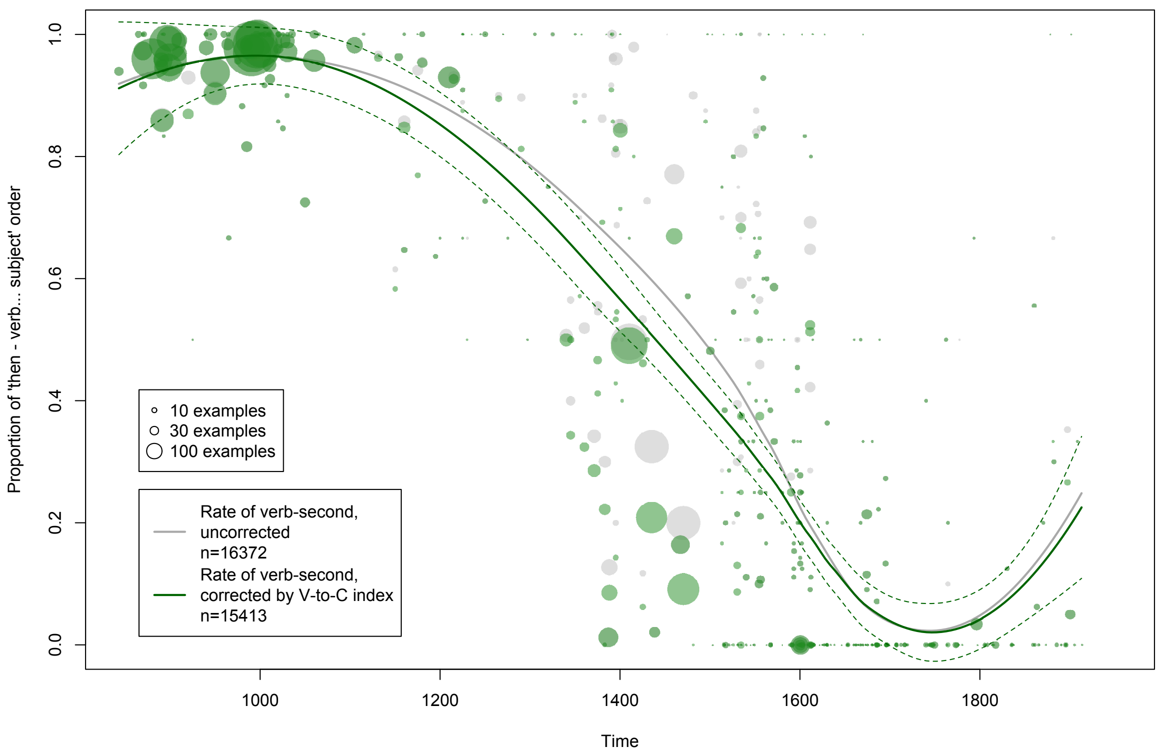



Figure 4.4: Development of verb-second order after adverbial *then* throughout the history of English

I am now in a position to evaluate the hypothesis under investigation here. For that purpose, I fitted a mixed-effects logistic regression models to the corrected data for high verb placement after *then*. It predicts the occurrence of verb-second after adverbial *then* from time and a random text effect. I considered the data only from the time horizon 1000–1500. The model is presented in table 4.5 below.

```
formula = VerbSecond ~  Year + (1 | Text)
          family = binomial, data = H1

Fixed effects:
                Estimate        Std.Error     z-value      p
Intercept       14.41           0.7361        19.57        <0.001***
Year            -0.0102412      0.0005656     -18.11       <0.001***

Random effect:
Text, N=117
Variance of random intercepts:  1.705

Null deviance:          6530 on 4870 degrees of freedom
Residual deviance:      3619 on 4868 degrees of freedom
AIC: 3625
```

Table 4.5: Logistic Regression Model 2: mixed-effects model for the decline in verb-second after *then*, 1000–1500

The model returns a 'Year' coefficient of -0.0102. At this rate of change, it would take 897 years for *then* to lose its verb-attracting property. The data comes from 117 text files, whose random intercepts show a variance of 1.705. This indicates substantial variation among texts. The model performs quite well. It fits significantly better than a null model (where the null model is an intercept only model) (Likelihood Ratio Test, $\chi^2$=3124.5, df=2, $p$<0.001) and presents a good fit to the data overall (Pseudo-$R^2_{marginal}$= 0.415, Pseudo-$R^2_{conditional}$=0.615). It has good classificatory power and categorizes considerably more cases correctly than the baseline (C-index= 0.918, classification accuracy= 83.5% vs. baseline: 60.6%).

The time course of the rise in subordinating *when* was already presented in section 4.3.2. The model changes only marginally if the difference between the two genre levels, prose and poetry, is ignored as a separate predictor.

One can now use the models fitted to the data sets for inversion after *then* and the realization of the temporal subordinator as *when* to find out which of these changes was likely actuated first. The answer to this question is the following. The model for main clause *then* returns the year 959 as the moment when the probability that this adverb would precede uninverted 'subject - verb' order first exceeded the 1%-threshold. In contrast, the model for subordinating *when* predicts that it was not before the year 1085 that there was a 1%-chance of finding this item as the formal realization of the subordinator in embedded temporal clauses. Therefore, the time courses of the two changes conform to the expectation of the first hypothesis - verb second order after *then* does indeed begin to decline earlier, by more than a century, than the onset of the change in subordinator form. Next, it is possible to use the same models to estimate how much the first change has progressed by the time the second change commences. The probability of finding the sentential modifier *then* followed by verb-second order in the year 1085, i.e., the inception point for the rise in subordinating *when*, is calculated as 96.4%. Appraising this percentage on a subjective basis, I would submit that it constitutes a very high value and hence indicates that the decline in verb-second order had advanced very little by the time *when* was innovated as a temporal subordinator. Hence, this result, too, is compatible with Hypothesis 1 - the loss of verb second order after *then* has not progressed so far relative to the onset of innovation in subordinator form as to make a causal association between them implausible.

The graph in figure 4.5 illustrates the above findings. It shows the decline in verb-second patterns after adverbial *then* in green and the rise of subordinating *when* in purple. Every text is represented by a data point and the logistic regression models for the two changes are shown by the fitted curves and their 95%-confidence intervals. One can see clearly that the green line for the loss of verb-second after *then* turns downwards before the purple line for the innovation in *wh*-based temporal subordination begins to rise. It is also quite evident that the first curve has declined only slightly by the time the second curve's ascent sets in.

In conclusion, the change in word order patterns after *then* commences before the rise of subordinating *when*. Yet, the former change does not become very widespread before the onset of the latter change. Hence, the evidence supports Hypothesis 1 and I will accept it for this reason.

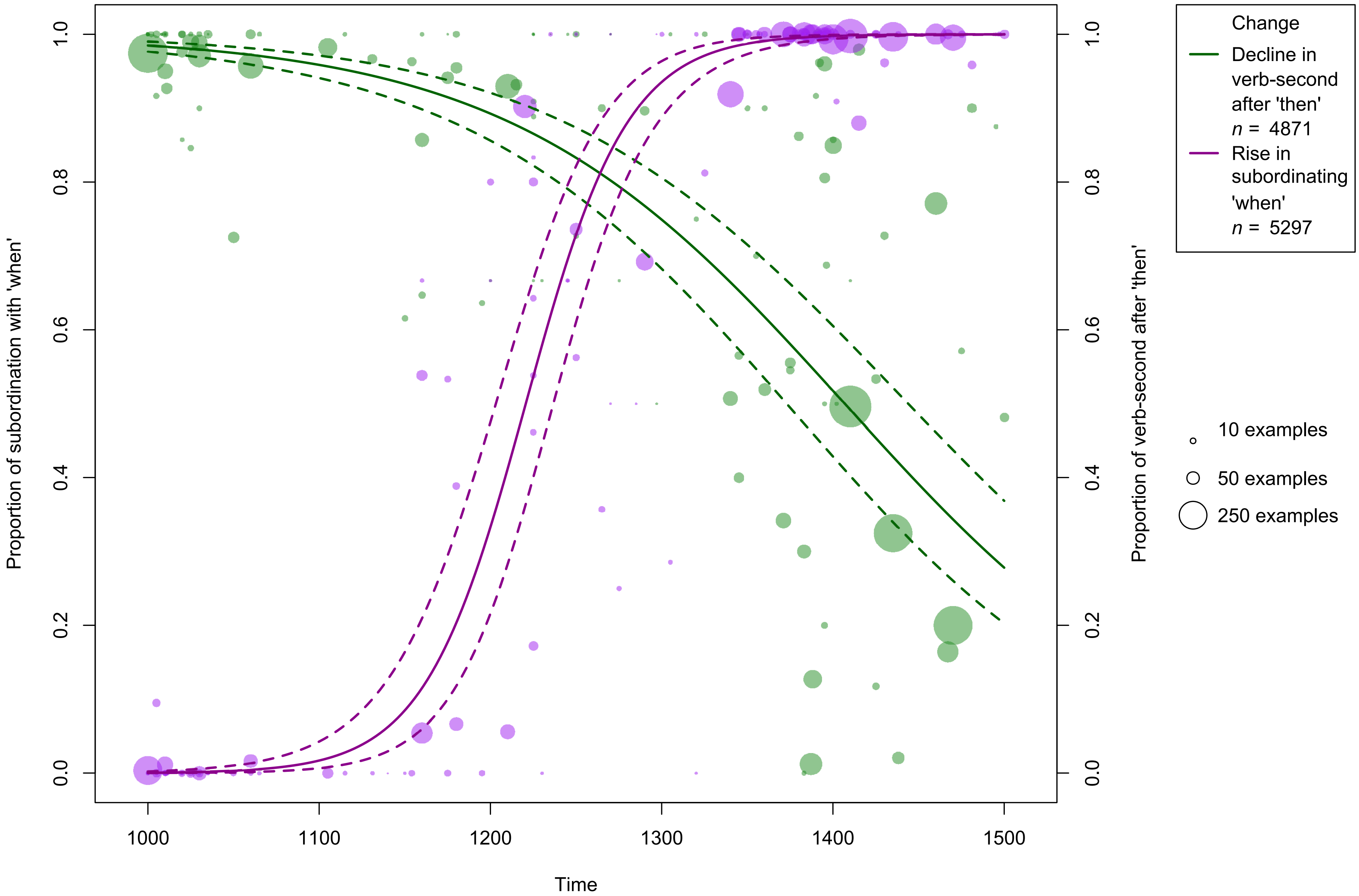



Figure 4.5: Time courses of the decline in verb-second after *then* and the rise of subordinating *when*

### 4.5.2 Hypothesis 2: Effect of Alternative Subordinating Strategies

Hypothesis 2 states that the probability of finding *when* as a subordinator should be greatly lowered in constructions that signal the dependent nature of the temporal clause through independent means. The prediction can be tested by using the presence or absence of an alternative subordinating strategy as a predictor for the formal realization of temporal subordinators.

#### 4.5.2.1 Data Collection

I coded all examples in the subordinator data set collected earlier (see section 4.3) for the presence or absence of an alternative subordinating strategy, i.e., for (i) correlative constructions, (ii) headed relativization or (iii) overt complementizers (see description in section 4.4.2.3). A headed temporal relative clause was annotated accordingly only if the adverb *then* functioned as its potential antecedent. The reason for this decision was (a) that the patterns *then then* and *then when* appear far more frequently than other potentially headed relatives and can therefore be regarded as more conventionalized expressions of subordination, and (b) that DPs or other adverbial constituents are both harder to retrieve from corpora and seem to me to function less rigorously as potential antecedents of temporal relative clauses. I did not code for the specific type of alternative subordinating strategy, but grouped all clauses that included an independent signal of subordination into one group. This choice was also justified by the fact that some clauses involved more than one indication of subordination. For example, the temporal clause in (64) involves *then* as the antecedent of a temporal relative but is also resumed by a second *then* in a correlative construction.

I ran search queries extracting instances of alternative subordinating strategies on the output files of the subordination data. The complement files of these queries must then contain all instances of temporal clauses without additional cues of embedding. Thus, my search queries separated the data into four logically possible groups. First, the subordinator could be expressed by a form of *then* and no other indication of subordination might be present, (63). Second, the temporal clause could be introduced by the same lexeme but simultaneously involve an alternative subordinating strategy, (64). Third, a temporal clause could be generated with initial *when* as the sole cue for its dependent nature, (65). Finally, a dependent clause could begin with the same *wh*-item but co-occur with an additional signal of embedding, (66). The indications of subordination are highlighted in boldface in the examples below.

(63) **subordinator *then***
**+ absence of alternative subordinating strategy**

hwar resteð þas mannes sawul [ **þone** se lychaman slepð ]
where rests the man's soul then the body sleeps

'Where will the soul of a man rest when his body sleeps?' (cosolsat1,Sol_I:41.1.146)

(64) **subordinator *then***
**+ presence of alternative subordinating strategy**

& [ **ða ða** he slep, ] **ða** genam he an rib of his sidan
and then then he slept then took he a rib of his side

'And when he was asleep, he took a rib out of his side' (cootest,Gen:2.21.109)

(65) **subordinator *when***
**+ absence of alternative subordinating strategy**

Bote muche schome þu þoledes. [ **hwen** þu þat neauer sunne dides; was taken as untreowe ]
but much shame you suffered when you that never sin did was taken as untrue

'But you suffered great shame when you, who never committed any sin, were taken as a liar' (WooingLord,279.191.125)

(66) **subordinator *when***
**+ presence of alternative subordinating strategy**

bote mare schome þu þoledes [ **hwen þat** te sunefule men i þi neb spitted ]
but more shame you suffered when that the sinful men in your face spat

'But you suffered more shame when those sinful men spat in your face' (WooingLord,279.197.131)

I did not conduct a search query for the Old English *when*-data as it only consisted of five examples. Instead, I evaluated those structures manually. None of them included an alternative subordinating strategy. This left 1 (Old English) + 2 (Middle English) = 3 search queries to be conducted.[8]

#### 4.5.2.2 Results

Out of 5,297 temporal subordinate clauses in my data set, 928 do and 4,369 do not involve an additional marker of subordination. I traced the frequency of alternative subordinating strategies in temporal subordinate *then*- and *when*-clauses through time from c. 1000 to 1500. The results can be summarized as follows. The subordinator *then* is more likely to co-occur with an additional signal of embedding than subordinating *when* overall ($\chi^2$=582.2, df=1, $p$<0.001, odds ratio = 5.70, 95%-confidence interval [4.89 – 6.65]) and throughout the era of their shared coexistence. In fact, some late Old English texts show the presence of an alternative subordinating strategy, especially correlatives, for more than 80% of all temporal *then*-clauses. In the 13$^{\text{th}}$ century, the proportion of *then*-clauses including an additional signal of subordination levels off at c. 20%. The development of *then*-clauses becomes difficult to track after c. 1350, when their frequency becomes vanishingly small overall. In contrast, the subordinator *when* rarely involves alternative subordinating strategies at first. While it is hard to estimate the exact proportion before the 12$^{\text{th}}$ century on account of the low number of relevant examples at that time, it seems fair to say that it does not exceed 10% until c. 1300. From the early 14$^{\text{th}}$ century on, there is a steady increase in the employment of additional cues for embedding, most importantly overt complementizers, reaching a peak in the middle of the 15$^{\text{th}}$ century at c. 20%, before the frequency begins to drop again.

Figure 4.6 illustrates the history of the proportion of alternative subordinating strategies in temporal *then*- and *when*-clauses. The graph presents the development in the form of data points for every text and smoothing Loess curves with their 95%-confidence intervals in orange for *then* and in turquoise for *when*.

Next, I fitted a logistic regression model to the data set. It predicts the realization of a temporal subordinator as *when* from time, genre, the presence or absence of an alternative subordinating strategy, this last variable's interaction with time, and a random text variable. The resulting model is shown in table 4.6.

```
formula = When ~  Year + Genre + AltSubStr + Year:AltSubStr + (1 | Text),
               family = binomial, data = H2
```

Fixed effects:

| | Estimate | Std.Error | z-value | $p$ |
|---|---|---|---|---|
| Intercept | -37.1486 | 3.0190 | -12.305 | <0.001*** |
| Year | 0.031038 | 0.002422 | 12.817 | <0.001*** |
| Genre(Prose→Poetry) | -1.2241 | 0.4955 | -2.471 | 0.01349* |
| AltSubStr(Absent→Present) | -9.7151 | 3.6510 | -2.661 | 0.00779** |
| Year:AltSubStr(Absent→Present) | 0.006736 | 0.002883 | 2.337 | 0.01946* |

```
Random effect:
Text, N=124
Variance of random intercepts:  3.164

Null deviance:         5952 on 5296 degrees of freedom
Residual deviance:     1222 on 5291 degrees of freedom
AIC: 1234
```

Table 4.6: Logistic Regression Model 3: mixed-effects model for alternative subordinating strategies

The model estimates an increase in the log-odds of *when* of 0.031 for every additional year that the change progresses. This value refers to temporal *when*-clauses without alternative subordinating strategies in prose texts. The 'Year' variable is a highly significant predictor in this model. Changing the genre variable from prose to poetry lowers the overall probability to find *when* as a subordinator by -1.2 log-odds. The genre variable is significant at the 5%-level. A total of 124 texts contributed to the data set whose random intercepts show a very substantial variance of 3.16. These findings replicate the effects of my earlier model for the rise of *when* as a temporal subordinator in section 4.3.

[8] The CorpusSearch queries, their complement file and summary excel sheet for the alternative subordinating strategy data can be found on the Dissertation DVD under `Chapter 4 - WhenThen/03 Hypothesis Testing/H2/`.

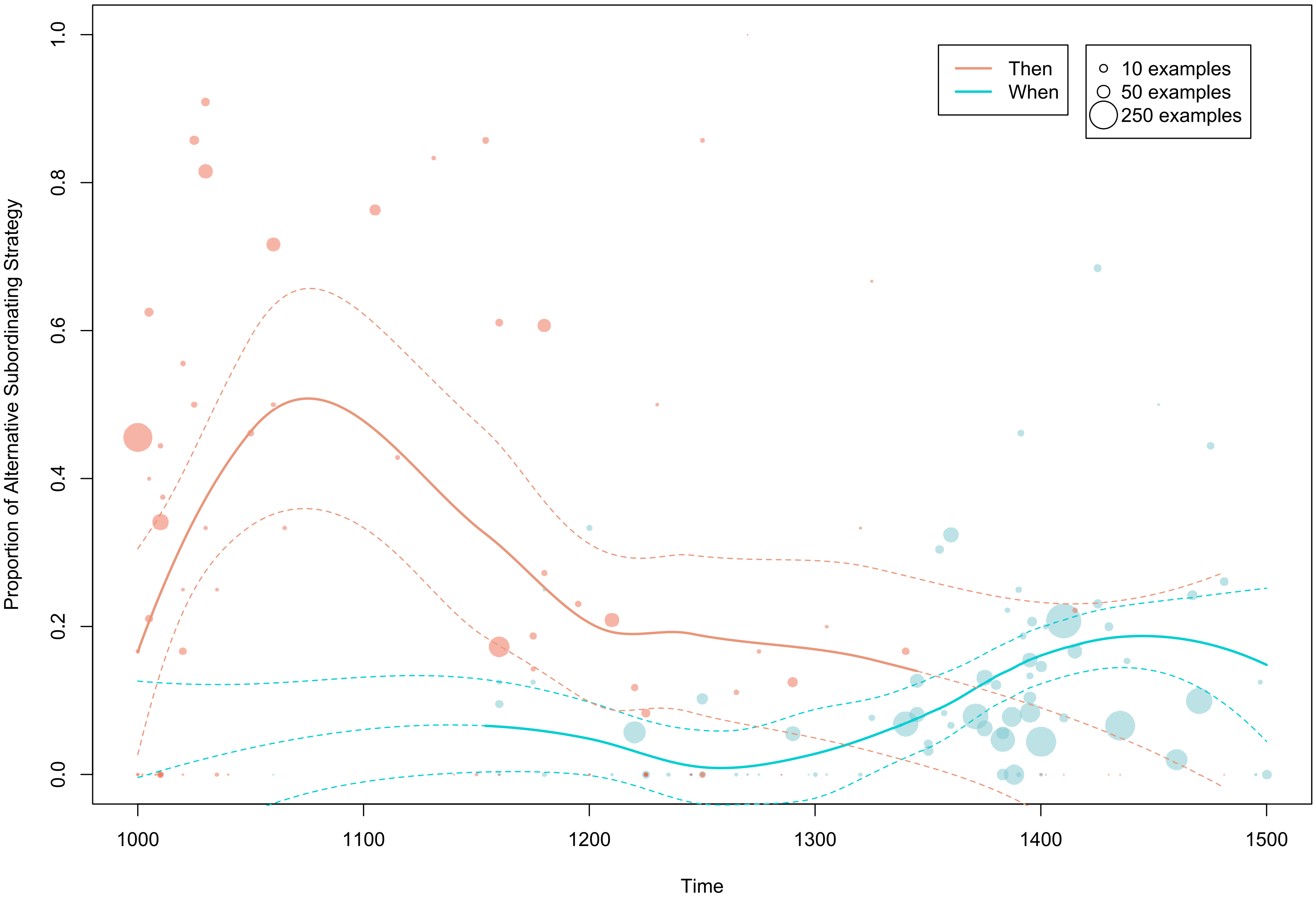



Figure 4.6: Development of frequency of alternative subordinating strategies in *then*- and *when*-clauses

The next line in table 4.6 lists the crucial variable for the investigation of the present hypothesis. It estimates the change in the probability of *wh*-based temporal subordination as the level of the variable 'alternative subordinating strategy,' abbreviated 'AltSubStr,' is changed from 'absent' to 'present.' If an additional signal of embedding is found in a temporal clause, its base probability of being introduced by the item *when* is reduced by -9.7 log-odds. The effect is statistically significant at the 1%-level in this model. It seems that dependent temporal *then*-clauses co-occurring with an independent signal of subordination do indeed seem to survive more robustly than their bare counterpart without such reinforcement. For example, collapsing the genre variable, one would predict that temporal clauses without an alternative subordinating strategy first have a 1%-chance of realizing their subordinator as *when* in the year 1069 whereas temporal clauses with such a strategy would not break through the same threshold before the year 1134. Therefore, this findings is precisely in accordance with my second hypothesis.

The final variable of the model estimates the difference in the rate of change as an alternative subordinating strategy is added to a temporal subordinate clause. Here, the activation of this condition is predicted to cause an acceleration in the rate of change to an increase of 0.0310 + 0.0067 = 0.0377 log-odds of subordinating *when* per year. The interaction term is significant at the 5%-level. The inclusion of the variable is justified by a likelihood ratio test on the difference in residual deviance between a model without the interaction term and the above model with the interaction term ($\chi^2$=6.34, df=1, $p$=0.0118). The steeper slope of the regression curve for examples with alternative subordinating strategies can be explained in the following way. Temporal *then*-clauses remain more likely with than without additional subordinating cues. However, they may fade out of the language in both contexts at the same time. Perhaps the simultaneous disappearance is due to the frequency of *then*-clauses dropping below some critical value in the majority context with no additional signals of subordination. Therefore, the relative frequency of conservative vis-à-vis innovative temporal subordination will shift more severely in the context with than the context without alternative subordinating strategies at the time when *then*-clauses eventually disappear. A logistic regression model will register this effect as a divergence in the rate of change. The statistically significant interaction between time and the presence or absence of additional cues for subordination in the above model can therefore be regarded as an ephemeral effect caused by one of the competing variants fading away in all contexts of use. It does not contradict the expectations expressed in Hypothesis 2.

The model performs well. It fits significantly better than a null model (where the null model is an intercept only model) (Likelihood Ratio Test, $\chi^2$=4881.7, df=5, $p$<0.001) and presents an adequate fit to the data (Pseudo-$R^2_{marginal}$= 0.794, Pseudo-$R^2_{conditional}$=0.895). It has very good classificatory power and classifies considerably more cases correctly than the baseline (C-index= 0.990, classification accuracy= 96.3% vs. baseline: 75.0%).

The data set for alternative subordinating strategies is illustrated in figure 4.7. The data points represent the frequency of *wh*-based temporal subordination for each text file. Cases without alternative subordinating strategies are shown in dark purple, occurrences with alternative subordinating strategies in light purple. The graph also presents the regression curves and their 95%-confidence intervals from two independent mixed-effects logistic regression models predicting the occurrence of temporal *when*-clauses in these two contexts in the same respective colors. One can clearly see that the light purple line indicating temporal clauses reinforced by a supplementary subordination strategy falls below the dark purple line representing ordinary temporal clauses without such a strategy. This highlights the fulfilled prediction encapsulated in Hypothesis 2 - temporal subordinators surface more frequently with the conservative variant *then* in the former than the latter context.

To conclude, the presence of an alternative subordinating strategy significantly increases the probability of encountering a conservative temporal *then*-clause. This finding supports the assertion of the second hypothesis so that one is justified in accepting it as well.

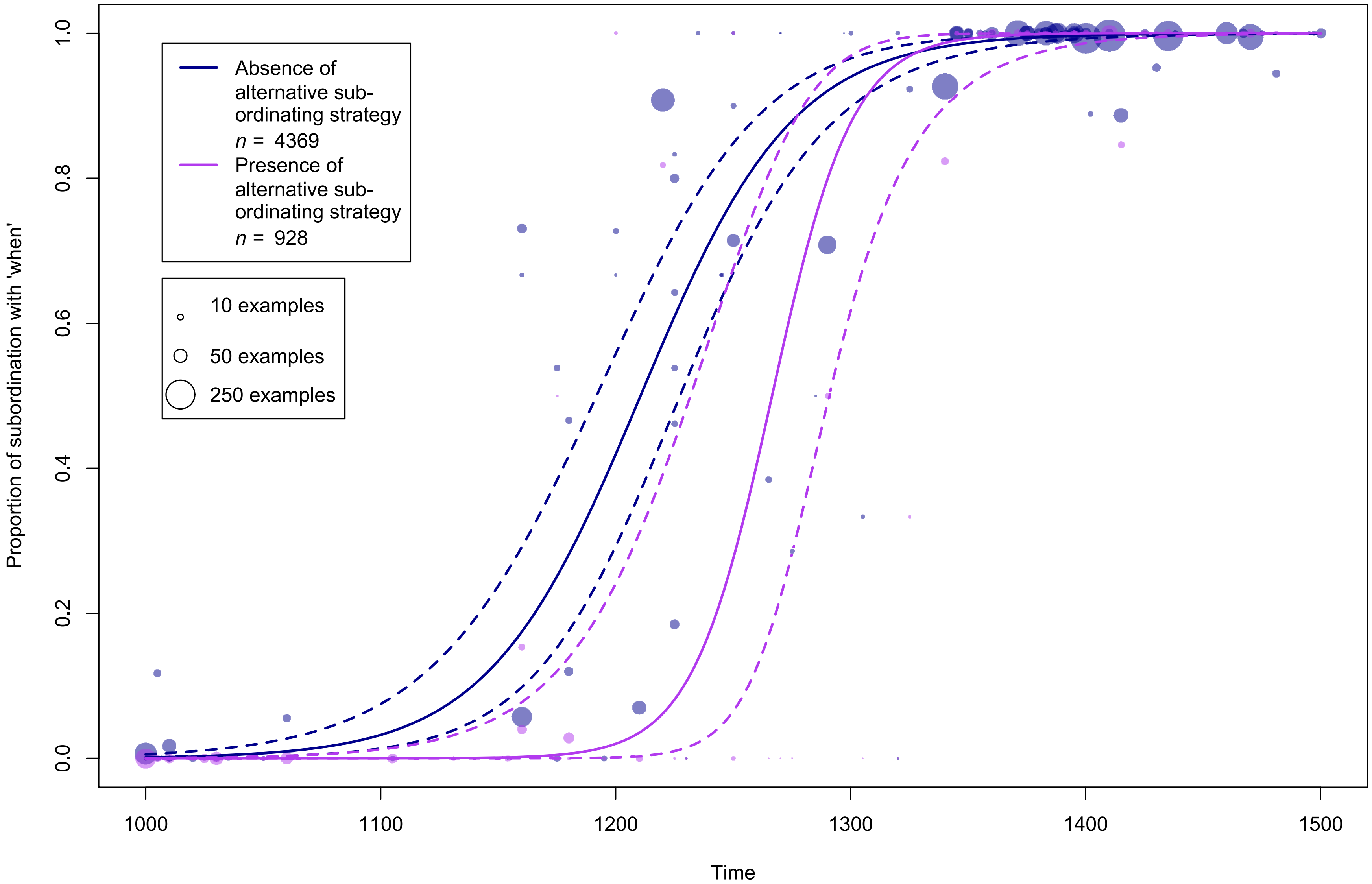



Figure 4.7: Rise in subordinating *when* by absence and presence of alternative subordinating strategies

### 4.5.3 Summary

In this part, I tested two hypotheses that were targeted at the causal component of an explanation for the Middle English rise in *wh*-based temporal subordination. This explanation states that the change occurred as a consequence of the loss of word-order patterns disambiguating the proper interpretation of the item *then*. I could provide evidence for both of them.

The investigation of Hypothesis 1 showed that the change in word order patterns after *then* was just beginning to obscure its proper interpretation by the time *when* was recruited as a subordinator. The relative timing of the two changes under consideration is thus compatible with a causal association between them. The proposed explanation for the development of subordinating *when* in terms of an internally triggered change does not, strictly speaking, receive positive corroboration from this study. However, the result does mean that the explanation has survived a realistic falsification attempt so that its correctness becomes somewhat more probable.

The hypothesis test for the second prediction revealed that the occurrence of a temporal subordinate clause introduced by *then* is significantly more likely if an alternative subordinating strategy is present than if it is not. This finding may offer some positive evidence for the correctness of the hypothesized causal influence of word order changes on the formal realization of temporal subordinators. In evolutionary terms, one could interpret the result to show that unambiguous identification of *then*'s grammatical role under the presence of independent cues for subordination lowers the advantage of innovative *when* over conservative *then* so that the intruding form cannot easily assert itself in that context. In more functional terms, it could be said that a supplementary marker of embedding for *then*-clauses would reduce the need that speakers feel to introduce a new, unambiguous subordinator. Either way, the fact that an additional signal of subordination seems to protect *then* from being ousted by *when* to a certain degree suggests that the unequivocal recognition of *then* as an adverb or as a subordinator plays an important role in the change. This, in turn, supports the assumption that the perturbation in the interpretability of *then* due to the loss of verb-second order after its function as a sentence adverb should exert a causal influence on the adoption of *when* as a new temporal subordinator.

## 4.6 Assessment

Middle English sees the rise in temporal subordination based on the *wh*-item *when*. The innovation displays the core characteristics of a common replacement change as defined in chapter 1: It involves two, formally identifiable variants, conservative *then* and innovative *when*, is attested from its earliest to its latest stage in the historical text record, the transitional period running from c. 1125–1375, and does not appear to be hampered by prescriptive pressures or other cultural contingencies. In fact, the change proceeds in such a straightforward way that I could measure it equally successfully in prose texts and a newly created poetry corpus (Zimmermann 2015). As a consequence, the replacement of subordinating *then* with *when* offered a fruitful test case for an approach to the holy grail of historical linguistics - the question of *why* a particular change occurred at the time that it did. I will now attempt to assess which general lessons about linguistic change can be learned from this study and in how far I have succeeded in my aim to provide an adequate explanation for the change.

My inquiry led me to formulate a couple of general thoughts, or perhaps just opinions, on the nature and scientific study of linguistic changes. They are the following.

(1) **How (not) to think about linguistic change**. One should be skeptical about claims positing a deterministic link between one necessary and sufficient cause and a subsequent linguistic innovation. There are two main reasons for this.

First, the cognitive, sociological and historical conditions on the expression of language are so complex that the conceptualization of linguistic change as the deduction of a consequent from one antecedent cause is almost certainly bound to be excessively unrealistic. Rather, one should entertain the possibility that linguistic innovations emerge from a criss-crossing network of interrelated and mutually influencing causes conspiring together to make a change more likely than not. Indeed, I could list a relatively large number of potential reasons for the rise of *when*, such as French influence, analogical extension based on other areas of the language system in which *wh*-items become dominant, and the loss of verb-second order after adverbial *then*. Several other possible reasons could probably be added (see e.g., Yamakawa 1969: fn. 69). It does not seem too far-fetched to assume that such a multi-factorial conception of determinants for linguistic changes can offer more comprehensive and adequate explanations in general and not just in the specific case investigated here.

Secondly, the formulation of reasons for linguistic changes in terms of direct 'if - then' statements seems to me to misconstrue the nature of such causes as unchanging, universal preference principles or economy conditions (e.g., Clark and Roberts 1993). I have argued here instead that one should compare linguistic entities to organisms whose change is best understood in terms of evolutionary dynamics. To give an analogy, a biologist might formulate the following principle of change, ‘a colder climate will cause a mammalian animal to develop thicker fur.’ While not unreasonable, the consequent of this statement does not necessarily follow for the species could also stay the way it is, migrate, find a new niche under ground or die out etc. Similarly, a cause for a linguistic change formulated as such a material implication, say ‘if a language loses verbal inflections, it will also lose V-to-I movement,’ will miss the point that the advantage of one grammatical option over another can only be assessed in relation to the specific linguistic environment in which it is used, even if several supporting cases can be adduced. In the test case analyzed in this chapter, I tried to illustrate such evolutionary reasoning by laying out a case for low adaptability of *then*-based subordination in a grammatical environment in which conditioning verb-second patterns after adverbial *then* are disappearing. The idea is that such a grammar will open up the opportunity for an innovative form to intrude on the linguistic space of its conservative competitor but does not make such a development inevitable.

(2) **How to test hypothesized causal associations**. It seems to be extremely difficult to derive testable predictions from an alleged link between some predictor and a linguistic outcome variable that specifically target the causal relation between the two dimensions. Often, it is hard enough to demonstrate just that such links can at least be regarded as plausible to begin with. Nevertheless, linguists should not be satisfied with mere demonstrations of plausibility but demand that assertions of supposed causes for a change be independently corroborated by hypothesis tests that could realistically falsify such claims. In the present study, I provided some limited evidence for my contention that word order changes contributed to the rise of subordinating *when* in the form of two hypothesis tests. The observed time course of the two potentially linked developments as well as an effect of alternative subordinating strategies are expected under my conceptualization of the change but may not be easy to account for otherwise or would have to be regarded as purely accidental.
In part, the deduction of falsifiable hypotheses from an explanation for a linguistic change may be due to the aforementioned complex, interwoven web of multiple causes for a change and the ever-changing nature of grammatical and cultural systems. Hence, it might be necessary to advance statistical techniques used to model linguistic changes. Perhaps more sophisticated multi-factorial analyses or structural equation modeling, commonly used in psychology, can be helpful in the future to assess the effects of unobserved, latent causes on linguistic innovations. Unfortunately, the application of such methods to the case at hand was not a feasible option. Thus, I limited my empirical studies to the influence of word order changes on the recruitment of *when* as a subordinator, thereby disregarding the effects of other plausible causes.

Given these substantial challenges for the investigation of the *why*-question of linguistic change, it should not be surprising that a large number of open questions remain with respect to the emergence of subordinating *when*. (i) How could one spell out more explicitly the supposed evolutionary dynamics bestowing on *when* and advantage over *then*, e.g., in terms of Yang’s (2002) or Jäger’s (2008) models of language change? (ii) Which testable predictions can be derived from other potential causes of the change, such as French calquing or analogy? In which ways could such factors be included in statistical models? How could one operationalize the degree of French influence on a text? How do other forms of *wh*-based subordination emerging during Middle English tie in with the developments reported here? (iii) Are there other testable predictions that could lend more credence to the conjecture that the development of *when* as a subordinator can in part be attributed to the loss of conditioning word order patterns? (iv) Are there potential confounding factors that should be controlled for, such as a temporal clause’s subject type, verb position, or information-structure? (v) Are there specific examples that should be considered individually because they represent atypical and idiosyncratic outliers?

Despite these open methodological and empirical questions, I believe that the rise in subordinating *when* provides one of the more convincing cases of a linguistic change whose cause has been examined in sufficient detail to become truly probable. Historical linguists, especially those working in a generative framework, have often assumed that one linguistic change could directly contribute to the emergence of a subsequent one. As far as it goes, however, the present study constitutes a rare example of such a supposedly system-internal motivation for language change that actually receives some empirical support. In this respect, it is my hope that the case study offered in this chapter will be a meaningful contribution to the study of causal determinants of linguistic changes.

# 5

# Conclusion

*"Endless forms most beautiful and most wonderful have been, and are being, evolved."*
– Charles Darwin *On the Origin of Species*, 1859

In the preface to her book *Case Marking and Reanalysis* (1995), Allen expresses her hope that studies on historical English syntax can contribute equally to an advancement in knowledge about general linguistic principles of syntactic change as well as about specific historical aspects of the evolution of the English language.

> [A]lthough my deepest interest is in learning more about the nature of human language in general, this does not mean that I have no interest in the history of the English language for its own sake, or that I view the history of this language solely as a means of learning more about language in general [...]. [I]t seems to me essential to bridge the gap between philology and theoretical linguistics. Linguists must rely on the fruits of the labours of philologists if they are to reconstruct an accurate picture of earlier stages of English. Conversely, the scholars involved in English Studies could profit from a better understanding of the findings of modern linguistics. (Allen 1995: vi)

I hope to have achieved a similar two-fold aim in this dissertation. In my study on late Modern American English possessive *have*, I provided solid support for the Constant Rate Hypothesis, one of the best corroborated abstract principles of language change, and presented an example of data management based on a modern, large corpus. At the same time, I reported several interesting features about the specific change itself, for instance, an effect of fictional vs. non-fictional genres, distinct preferences for certain adjunction sites by different adverbs, and, of course, the exact time course of the development. The second case study on Early English conjunct clauses showed that syntactic changes can interact with each other in a way that is traceable in corpora and used predicted violations of constant rate effects to demonstrate the existence of such interactions. Simultaneously, I offered a parsimonious explanation of the special status of conjunct clauses in Old English and produced many individual examples to argue in substantial detail for particular analyses of common, macro-structural constructions of that language stage. Finally, my third investigation, dealing with the Middle English replacement of subordinating *then* with *when*, allowed me to reflect on the conceptualization of and good methodologies for the *why*-question of linguistic innovations in general, but also to discover specific facets of the change, such as a constant rate effect between prose and poetry or the exact delimitation of the transitional period, in particular. I am therefore optimistic that my thesis could further the knowledge of syntactic changes relevant to both the theoretical linguist and the anglicist scholar.

The general methodology of studying syntactic change used in this thesis, while perhaps not particularly innovative, proved to be a successful pathway to relatively certain insights and corroboration. It involved (i) building formal grammar models and (ii) testing them quantitatively in corpus studies. It is not always easy to apply this technique in a useful way, i.e., to derive hypotheses that are both quantitatively testable and yet have a bearing on analyses embodied in formal grammars. Yet, both theoretical guidance and empirical justification have been regarded as essential aspects of the scientific process since the enlightenment. The point of view is well epitomized in Kant's famous statement, "Concepts without percepts are empty, percepts without concepts are blind" (*Critique of Pure Reason*, B75). Neither aspect should therefore simply be left out of a study. For the cases I discussed, theoretical concepts, such as grammar competition, phrase structure rules or adjunction sites, would have remained meaningless if they had not been applied empirically to specific changes, different grammatical contexts, or studies on correlations between adverb position and rates of change etc. Conversely, it would have been impossible to decide what units compete with each other, which contexts to collect data for, or what predictions to test, had an explicit grammar model not guided every step along the analysis process. I would therefore like to think that I have led by example, and could illustrate a systematic and profitable way to unite formal and quantitative approaches to the study of syntactic changes.

I would like to close my thesis like I started it - with an analogy from biology. Evolution by natural selection accounts for all diversity of life, from the tiniest microbes, to the mightiest vertebrates, for, in Darwin's words, "endless forms, most beautiful and most wonderful." So, too, if, as most linguists assume, the faculty to use language is a genetically determined, universal trait of humans, all diversity among human languages must be a product of language-internal and external pressures to adapt. Every last aspect of the wondrous and unfathomable variety of language, from accusative to absolutive, from clicks to case stacking, from valency diatheses to verb-first, arises from, and indeed can only arise from, gradual changes accumulating over enormous periods of time. I stand in awe of this realization and wonder what it may yet have to teach us.

# Bibliography

Adger, David (2003) *Core Syntax: A Minimalist Approach.* Oxford: Oxford University Press.

Aelbrecht, Lobke (2009) *You have the Right to Remain Silent: The Syntactic Licensing Conditions on Ellipsis.* PhD Dissertation, Catholic University of Brussels.

Allen, Cynthia L. (1995) *Case Marking and Reanalysis: Grammatical Relations from Old to Early Modern English.* Oxford: Oxford University Press.

Altmann, Gabriel (1983) 'Das Piotrowski-Gesetz und seine Verallgemeinerungen.' In: Best, Karl-Heinz and Jörg Kohlhase (eds.) *Exakte Sprachwandelforschung. Theoretische Beitrge, statistische Analysen und Arbeitsberichte.* Göttinger Schriften zur Sprach- und Literaturwissenschaft. Bd. 2. Göttingen: Herodot, pp. 54–90.

Andor, József (2013) 'Interrogating Possessive *have*: A Case Study.' *Argumentum* **9**, 99–107.

Andrew, Samuel O. (1940) *Syntax and Style in Old English.* Cambridge: Cambridge University Press.

Andrews, Avery (2007) 'Input and Glue in OT-LFG.' In: Zaenen, Annie, Jane Simpson, Tracy Holloway King, Jane Grimshaw and Joan Maling (eds.) *Architectures, Rules, and Preferences: Variations on themes by Joan W. Bresnan.* Stanford, CA: CSLI Press, pp. 319–340.

Asudeh, Ash (2012) *The Logic of Pronominal Resumption.* Oxford: Oxford University Press.

Asudeh, Ash, Gianluca Giorgolo and Ida Toivonen (2014) 'Meaning and Valency.' *Proceedings of the LFG Conference* **19**, 68–88.

Asudeh, Ash, Mary Dalrymple and Ida Toivonen (2013) 'Constructions with Lexical Integrity.' *Journal of Language Modelling* **1.1**, 1–54.

Asudeh, Ash and Richard Crouch (2002) 'Derivational Parallelism and Ellipsis Parallelism.' *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL)* **21**, 1–14.

Auwera, Johan van der, Nina Dobrushina and Valentin Goussev (2005) 'Imperative-Hortative Systems.' In: Haspelamth, Martin, Matthew S. Dryer, David Gil and Bernard Comrie (eds.) *The World Atlas of Language Structures.* Oxford: Oxford University Press, pp. 296–299.

Axel, Katrin (2007) *Studies on Old High German Syntax: Left Sentence Periphery, Verb Placement and Verb-Second.* Amsterdam: John Benjamins.

Baker, Mark (2001) *The Atoms of Language: The Mind's Hidden Rules of Grammar.* New York: Basic Books.

Bates, Douglas, John Chambers, Peter Dalgaard, Robert Gentleman, Kurt Hornik, Ross Ihaka, Tomas Kalibera, Michael Lawrence, Friedrich Leisch, Uwe Ligges, Thomas Lumley, Martin Maechler, Martin Morgan, Duncan Murdoch, Paul Murrell, Martyn Plummer, Brian Ripley, Deepayan Sarkar, Duncan Temple-Lang, Luke Tierney and Simon Urbanek (2008) *R: A Language and Environment for Statistical Computing* `http://www.R-project.org/` (Accessed September 2016) Vienna: R Foundation for Statistical Computing.

Bean, Marian C. (1983) *The Development of Word Order Patterns in Old English.* London: Croom Helm.

Bech, Kristin (2001) *Word Order Patterns in Old and Middle English: A Syntactic and Pragmatic Study.* PhD Dissertation, University of Bergen.

Bech, Kristin (2015) 'Old 'Truths', New Corpora: Revisiting the Word Order of Conjunct Clauses in Old English.' *English Language and Linguistics* **21.1**, 1–25.

Beesley, Kenneth R. and Lauri Karttunen (2003) *Finite State Morphology.* Stanford, CA: CSLI Press.

Bellert, Irena (1977) 'On semantic and distributional properties of sentential adverbs.' *Linguistic Inquiry* **8.2**, 337–351.

Bender, Emily M. and Ivan A. Sag (2000) 'Incorporating Contracted Auxiliaries in English.' In: Cann, Ronnie, Claire Grover and Philip Miller (eds.) *Grammatical Interfaces in HPSG.* Stanford, CA: CSLI Press, pp. 17–32.

Bergen, Linda van (2003) *Pronouns and Word Order in Old English: With Particular Reference to the Indefinite Pronoun Man.* New York: Routledge.

Besten, Hans den (1983) 'On the Interaction of Root Transformations and Lexical Deletive Rules.' In: *On the Formal Syntax of the Westgermania.* Amsterdam: John Benjamins, pp. 47–131.

Biberauer, Theresa and Ian G. Roberts (2005) 'Changing EPP-Parameters in the History of English: Accounting for Variation and Change.' *English Language and Linguistics* **9**, 5–46.

Bloomfield, Leonard (1942) *Outline Guide for the Practical Study of Foreign Languages.* Baltimore: Linguistic Society of America at the Waverly Press.

Bobaljik, Jonathan D. (1995) *Morphosyntax: The Syntax of Verbal Inflection.* PhD Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Bobaljik, Jonathan D. (2003) 'Floating Quantifiers: Handle with Care.' In: Cheng, Lisa and Rint Sybesma (eds.) *The Second Glot International State-of-the-Article Book.* Berlin: de Gruyter, pp. 107–148.

Borsley, Robert D. and Frederick J. Newmeyer (2009) 'On Subject-Auxiliary Inversion and the Notion"Purely Formal Generalization".' *Cognitive Linguistics* **20.1**, 135–143.

Boškovič, Željko and Jairo Nunes (2007) 'The Copy Theory of Movement: A view from PF.' In: Corver, Norbert and Jairo Nunes (eds.) *The Copy Theory of Movement.* Amsterdam: John Benjamins, pp. 13–74.

Bresnan, Joan (1982) 'Control and Complementation.' In: Bresnan, Joan (ed.) *The Mental Representation of Grammatical Relations.* Cambridge, MA: MIT Press, pp. 282–390.

Bresnan, Joan (2001) *Lexical-Functional Syntax.* Oxford: Blackwell.

Bresnan, Joan and Jane Grimshaw (1978) 'The Syntax of Free Relatives in English.' *Linguistic Inquiry* **9.3**, 331–391.

Bruening, Benjamin (2016) *Old English Verb-Second-ish in a Typology of Verb-Second.* Manuscript, University of Delaware.

Bulkley Cobb, Caroline (2006) *The Syntax of Adverbs: An LFG Approach.* MA Dissertation, University of Oxford.

Büring, Daniel (2004) 'Negative Inversion.' *Proceedings of the North East Linguistic Society* **35**, 1–20.

Burnett, Heather and Elizabeth Ferch (2005) *The Loss of V2 Clause Structure in the History of English: A Review of the Literature.* Manuscript, Queen's University.

Butt, Miriam, Tracy Holloway King, María-Eugenia Niño and Frédérique Segond (1999) *A Grammar Writer's Cookbook.* Stanford, CA: CSLI Press.

Butters, Ronald R. (2001) 'Grammatical Structure.' In: Algeo, John (ed.) *The Cambridge History of the English Language.* Cambridge: Cambridge University Press, pp. 325–339.

Bybee, Joan L. (2012) 'Usage-Based Theory and Grammaticalization.' In: Narrog, Heiko and Bernd Heine (eds.) *The Oxford Handbook of Grammaticalization.* Oxford: Oxford University Press, pp. 69–78.

Campbell, Alistair (1970) 'Verse Influences in Old English Prose.' In: Rosier, James L. (ed.) *Philological Essays: Studies in Old and Middle English Language and Literature in Honour of Herbert Dean Meritt.* The Hague: Mouton, pp. 93–98.

Canale, Michael W. (1978) *Word Order Change in Old English: Base Reanalysis in Generative Grammar.* PhD Dissertation, McGill University.

Cardinaletti, Anna and Ian G. Roberts (1991) *Clause Structure and X-Second.* Manuscript, University of Venice and University of Geneva.

Carr, Charles T. (1933) 'Some Old High German Conjunctions.' *The Journal of English and Germanic Philology* **32.4**, 488–503.

Cheshire, Jenny (2003) 'Social Dimensions of Syntactic Variation: The Case of *when* Clauses.' In: Britain, David and Jenny Cheshire (eds.) *Social Dialectology: In Honour of Peter Trudgill.* Amsterdam: John Benjamins, pp. 245–261.

Chomsky, Noam (1986) *Barriers.* Cambridge, MA: MIT Press.

Chomsky, Noam (1991) 'Some Notes on Economy of Derivation and Representation.' In: Freiden, Robert (ed.) *Principles and Parameters in Comparative Grammar.* Cambridge, MA: MIT Press, pp. 53–82.

Chomsky, Noam (1995) *The Minimalist Program.* Cambridge, MA: MIT Press.

Chung, Sandra (2006) 'Sluicing and the Lexicon: The Point of No Return.' In: Cover, Rebecca T. and Yuni Kim (eds.) *Proceedings of the Annual Meeting of the Berkeley Linguistics Society 31.* Berkeley, CA: Berkeley Linguistics Society, pp. 73–91.

Cichosz, Anna (2010) *The Influence of Text Type on Word Order of Old Germanic Languages: A Corpus-Based Contrastive Study of Old English and Old High German.* Studies in English Medieval Language and Literature 27. Frankfurt on the Main: Peter Lang.

Cinque, Guglielmo (1999) *Adverbs and Functional Heads: A Cross-Linguistic Perspective.* Oxford: Oxford University Press.

Clark, Brady Z. (2004) *A Stochastic Optimality Theory Approach to Language Change.* PhD Dissertation, Stanford, CA.

Clark, Robin and Ian G. Roberts (1993) 'A Computational Model of Language Learnability and Language Change.' *Linguistic Inquiry* **24**, 299–345.

Corley, Kerry (2014) *The Constant Rate Hypothesis in Syntactic Change: Empirical Fact or 'Lies, Damned Lies, and Statistics'?.* BA Dissertation, University of Cambridge.

Craenenbroeck, Jeroen van and John Merchant (2013) 'Ellipsis Phenomena.' In: Dicken, Marcel den (ed.) *The Cambridge Handbook of Generative Syntax.* Cambridge: Cambridge University Press, pp. 701–745.

Croft, William (ed.) (2000) *Explaining Language Change: An Evolutionary Approach.* London: Longman Linguistics Library.

Crouch, Richard (1999) 'Ellipsis and Glue Languages.' In: Lappin, Shalom and Elabbas Benmamou (eds.) *Fragments: Studies in Ellipsis and Gapping.* Oxford: Oxford University Press, pp. 32–67.

Crouch, Richard, Mary Dalrymple, Ron Kaplan, Tracy King, John Maxwell and Paula Newman (1993) *XLE Documentation* `http://www2.parc.com/isl/groups/nltt/xle/doc/xle_toc.html/` (Accessed September 2016). Palo Alto Research Center (PARC).

Dalrymple, Mary (2001) *Lexical Functional Grammar.* Syntax and Semantics 34. New York: Academic Press.

Dalrymple, Mary and Irina Nikolaeva (2011) *Objects and Information Structure.* Cambridge: Cambridge University Press.

Dalrymple, Mary, Stuart M. Shieber and Fernando C. N. Pereira (1991) 'Ellipsis and Higher-Order Unification.' *Linguistics and Philosophy* **14.4**, 399–452.

Davies, Mark (2010) *The Corpus of Historical American English: 400 million words, 1810-2009.* `http://corpus.byu.edu/coha/` (Accessed 10 September 2013).

Davies, Mark (2012) 'Expanding Horizons in Historical Linguistics with the 400 Million Word Corpus of Historical American English.' *Corpora* **7**, 121–157.

Davis, Graeme (2006) *Comparative Syntax of Old English and Old Icelandic: Linguistic, Literary and Historical Implications.* Oxford: Peter Lang.

Declerck, Renaat H. C (ed.) (1997) *When-Clauses and Temporal Structure.* London: Routledge.

Delfitto, Denis (2006) 'Adverb Classes and Adverb Placement.' In: Everaert, Martin, Henk van Riemsdijk, Rob Goedemans and Bart Hollebrandse (eds.) *The Blackwell Companion to Syntax.* Volume 1. Malden, MA: Blackwell, pp. 83–120.

Denison, David (2001) 'Gradience and Linguistic Change.' In: Brinton, Laurel J. (ed.) *Historical Linguistics 1999: Selected Papers from the 14th International Conference on Historical Linguistics.* Amsterdam: John Benjamins, pp. 119–144.

Denison, David (2003) 'Log(ist)ic and Simplistic S-Curves.' In: Hickey, Raymond (ed.) *Motives for Language Change.* Cambridge: Cambridge University Press, pp. 54–70.

Devine, Andrew M. and Laurence D. Stephens (2006) *Latin Word Order: Structured Meaning and Information.* Oxford: Oxford University Press.

Dobzhansky, Theodosius (1973) 'Nothing in Biology Makes Sense except in the Light of Evolution.' *The American Biology Teacher* **35.3**, 125–129.

Dowty, David R. and Belinda Brodi (1984) 'The Semantics of 'Floated' Quantifiers in a Transformationless Grammar.' In: Cobler, Mark, Susannah MacKaye and Michael T. Westcoat (eds.) *Proceedings of the 3rd West Coast Conference on Formal Linguistics (WCCFL 3).* Stanford, CA: Stanford Linguistics Association, pp. 75–90.

Dreschler, Gea (2015) *Passives and the Loss of Verb Second.* Utrecht: LOT.

Dryden, Henry E. (1843) *The Art of Hunting by William Twici Huntsman to King Edward the Second.* Daventry: Thomas Barrett.

Dryer, Matthew S. (1986) 'Primary Objects, Secondary Objects and Antidative.' *Language* **62.4**, 808–845.

Ecay, Aaron W. (2015) *A Multi-Step Analysis of the Evolution of English* do-*Support.* PhD Dissertation, University of Pennsylvania.

Edelstein, Elspeth C. (2012) *The Syntax of Adverb Distribution.* PhD Dissertation, University of Edinburgh.

Ellegård, Alvar (1953) *The Auxiliary do: The Establishment and Regulation of its Use in English.*. Gothenburg Studies in English Stockholm: Almqvist and Wiksell.

Embick, David and Rolf Noyer (2001) 'Movement Operations after Syntax.' *Linguistic Inquiry* **32.4**, 555–595.

Engels, Eva (2010) 'Adverb Placement in English VP Ellipsis Constructions.' In: Götzsche, Hans (ed.) *Memory, Mind and Language.* Cambridge: Cambridge Scholars Publishing, pp. 149–162.

Engels, Eva (2012) *Optimizing Adverb Positions.* Linguistics Today No. 181. Amsterdam: John Benjamins.

Ernst, Thomas (2002) *The Syntax of Adjuncts.* Cambridge: Cambridge University Press.

Falk, Yehuda N. (2008) 'Functional Relations in the English Auxiliary System.' *Linguistics* **46.5**, 861–889.

Fiengo, Robert and Robert May (1994) *Indices and Identity.* Cambridge, MA: MIT Press.

Fillmore, Charles J. (1968) 'The Case for Case.' In: Bach, Emmons and Robet T. Harms (eds.) *Universals in Linguistic Theory.* New York: Holt, Rinehart, and Winston, pp. 1–89.

Fischer, Olga, Ans van Kemenade, Willem Koopman and Wim van der Wurff (2000) *The Syntax of Early English.* Cambridge: Cambridge University Press.

Foster, Robert (1975) 'The Use of *þa* in Old and Middle English Narratives.' *Neuphilologische Mitteilungen* **76**, 404–414.

Fourquet, Jean (1938) *L'Ordre des Elements de la Phrase en Germanique Ancien.* Strasbourg: Publications de la Faculté des Lettres de Strasbourg.

Fox, Danny J. (2000) *Economy and Semantic Interpretation.* Cambridge, MA: MIT Press.

Fox, Danny J. and Howard Lasnik (2003) 'Successive-Cyclic Movement and Island Repair: The Difference between Sluicing and VP-Ellipsis.' *Linguistic Inquiry* **34.1**, 143–154.

Frank, Anette and Annie Zaenen (2002) 'Tense in LFG: Syntax and Morphology.' In: Kamp, Hans and Uwe Reyle (eds.) *How we Say WHEN it Happens. Contributions to the Theory of Temporal Reference in Natural Language.* Tübingen: Niemeyer, pp. 17–51.

Frey, Werner and Karin Pittner (1999) 'Adverbialpositionen im Deutsch-Englischen Vergleich.' In: Doherty, Monika (ed.) *Sprachspezifische Aspekte der Informationsverteilung.* Berlin: Akademischer Verlag, pp. 14–40.

Fulk, Robert D. (1992) *A History of Old English Meter.* Philadelphia: University of Pennsylvania Press.

Fuss, Eric (2003) 'On the Historical Core of V2 in Germanic.' *Nordic Journal of Linguistics* **26.2**, 195–231.

Fuss, Eric (2017) 'Language Change.' In: Roberts, Ian G. (ed.) *The Oxford Handbook of Universal Grammar.* Oxford: Oxford University Press, pp. 459–485.

Fuss, Eric and Carola Trips (2002) 'Variation and change in Old and Middle English – on the validity of the Double Base Hypothesis.' *Journal of Comparative Germanic Linguistics* **4**, 171–224.

Geis, Michael L. (1970) *Adverbial Subordinate Clauses in English'.* PhD Dissertation, Cambridge, MA: MIT Press.

Gelderen, Elly van (2004) *Grammaticalization as Economy.* Amsterdam: John Benjamins.

Gelderen, Elly van (2013) 'Remarks and Replies: Null Subjects in Old English.' *Linguistic Inquiry* **44.2**, 271–285.

Gelderen, Elly van (2017) *The Main and Embedded Clauses in the History of English: Changes in Assertive and Non-Assertive Complements.* Manuscript, Arizona State University.

Gelman, Andrew (2005) 'Analysis of Variance: Why it is More Important than Ever.' *The Annals of Statistics* **33.1**, 1–53.

Gelman, Andrew and Jennifer Hill (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Analytical Methods for Social Research Cambridge: Cambridge University Press.

Goldberg, Adele E. (2003) 'Constructions: A New Theoretical Approach to Language.' *Trends in Cognitive Science* **7.5**, 219–224.

Goldberg, Adele E. (2006) *Constructions at Work: The Nature of Generalizations in Language.* Oxford: Oxford University Press.

Goldberg, Adele E. (2009) 'The Nature of Generalization in Language.' *Cognitive Linguistics* **20.1**, 93–127.

González-Diaz, Victorina (2008) *English Adjective Comparison: A Historical Perspective.* Amsterdam Studies in the Theory and History of Linguistic Science Amsterdam: John Benjamins.

Gries, Stefan Th. (2015) 'The Most Underused Statistical Method in Corpus Linguistics: Multi-level (and Mixed-effects) Models.' *Corpora* **10.1**, 95–125.

Grimshaw, Jane (1977) *English Wh-Constructions and the Theory of Grammar.* PhD Dissertation, Amherst: University of Massachusetts.

Haeberli, Eric (2000) 'Adjuncts and the Syntax of Subjects in Old and Middle English.' In: Pintzuk et al. (2000) , pp. 109–131.

Haeberli, Eric (2001) 'Speculations on the Syntax of Subordinate Clauses in Old English.' *Reading Working Papers in Linguistics* **5**, 201–229.

Haeberli, Eric (2002*a*) 'Inflectional Morphology and the Loss of Verb Second in English.' In: Lightfoot, David (ed.) *Syntactic Effects of Morphological Change.* Oxford: Oxford University Press, pp. 88–106.

Haeberli, Eric (2002*b*) 'Observations on the Loss of Verb Second in the History of English.' In: Zwart, Jan-Wouter and Werner Abraham (eds.) *Studies in Comparative Germanic Syntax: Proceedings from the 15th Workshop on Comparative Germanic Syntax (Groningen, May 26-27, 2000).* Amsterdam: John Benjamins, pp. 245–272.

Haeberli, Eric (2005) 'Clause Type Asymmetries in Old English and the Syntax of Verb Movement.' In: Batllori, Montserrat, Maria-Lluïsa Hernanz, Carmen Picallo and Francesc Roca (eds.) *Grammaticalization and Parametric Change.* Oxford: Oxford University Press, pp. 267–283.

Haeberli, Eric (2007) 'The Development of Subject-Verb Inversion in Middle English and the Role of Language Contact.' *Generative Grammar in Geneva* **5**, 15–33.

Haeberli, Eric (2010) 'Investigating Anglo-Norman Influence on Late Middle English Syntax.' In: Ingham (2010*b*), pp. 143–163.

Haeberli, Eric and Liliane Haegeman (1995) 'Clause Structure in Old English: Evidence from Negative Concord.' *Journal of Linguistics* **31.1**, 81–108.

Haeberli, Eric and Richard Ingham (2007) 'The Position of Negation and Adverbs in Early Middle English.' *Lingua* **117.1**, 1–25.

Haeberli, Eric and Tabea Ihsane (2016) 'Revisiting the Loss of Verb Movement in the History of English.' *Natural Language & Linguistic Theory* **34.2**, 497–542.

Haegeman, Liliane (1995) *The Syntax of Negation.* Cambridge: Cambridge University Press.

Haegeman, Liliane (2010) 'The Internal Syntax of Adverbial Clauses.' *Lingua* **120.3**, 628–648.

Haegeman, Liliane (2012) *Adverbial Clauses, Main Clause Phenomena, and the Composition of the Left Periphery.* Oxford: Oxford University Press.

Haegeman, Liliane and Jacqueline Guéron (1999) *English Grammar: A Generative Perspective.* Oxford: Blackwell.

Halverson, Per-Kristian (1983) 'Semantics for Lexical Functional Grammar.' *Linguistic Inquiry* **14.4**, 567–615.

Hardt, Daniel (1997) 'An Empirical Approach to VP Ellipsis.' *Computational Linguistics* **23.4**, 525–541.

Haumann, Dagmar (2007) *Adverb Licensing and Clause Structure in English.* Amsterdam: John Benjamins.

Hawkins, John (2000) 'The Relative Order of Prepositional Phrases in English: Going beyond Manner-Place-Time.' *Language Variation and Change* **11**, 231–266.

Hawkins, John A. (1994) *A Performance Theory of Order and Constituency.* Cambridge: Cambridge University Press.

Heny, Frank (1973) 'Sentence and Predicate Modifiers in English.' In: Kimball, John P. (ed.) *Syntax and Semantics.* Vol. 2 New York: Academic Press, pp. 217–245.

Hilpert, Martin (2013) *Constructional Change in English: Developments in Allomorphy, Word Formation and Syntax.* Cambridge: Cambridge University Press.

Hilpert, Martin and Stefan Th. Gries (2010) 'Modeling Diachronic Change in the Third Person Singular: a Multifactorial, Verb- and Author-specific Exploratory Approach.' *English Language and Linguistics* **14.3**, 293–320.

Hinskens, Frans (1998) 'Variation Studies in Dialectology and Three Types of Sound Change.' *Sociolinguistica* **12.1**, 155–193.

Hinterhölzl, Roland and Svetlana Petrova (2010) 'From V1 to V2 in West Germanic.' *Lingua* **120.2**, 315–328.

Hobbs, Jerry R. and Andrew Kehler (1997) 'A Theory of Parallelism and the Case of VP Ellipsis.' *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, pp. 394–401.

Huddleston, Rodney (1976) 'Some Theoretical Issues in the Description of the English Verb.' *Lingua* **40.4**, 331–383.

Huddleston, Rodney and Geoffrey K. Pullum (2002) *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press.

Hudson, Richard A. (1980) *Sociolinguistics.* Cambridge: Cambridge University Press.

Ingham, Richard (2010*a*) 'The Transmission of Anglo-Norman: Some Syntactic Evidence.' in *The Anglo-Norman Language and its Context.* Ingham (2010*b*), pp. 164–182.

Ingham, Richard (ed.) (2010*b*) *The Anglo-Norman Language and its Context.* York: York Medieval Press.

Jackendoff, Ray (1972) *Semantic Interpretation in Generative Grammar.* Cambridge, MA: MIT Press.

Jäger, Gerhard (2004) *Evolutionary Game Theory for Linguists. A Primer.* Manuscript, Stanford University and University of Potsdam.

Jäger, Gerhard (2008) 'Applications of Game Theory in Linguistics.' *Language and Linguistics Compass* **2**, 408–421.

Janda, Richard D. and Brian D. Joseph (2003*a*) 'On Language, Change, and Language Change - Or of History, Linguistics and Historical Linguistics.' in *The Handbook of Historical Linguistics.* Janda and Joseph (2003*b*), pp. 3–180.

Janda, Richard D. and Brian D. Joseph (eds.) (2003*b*) *The Handbook of Historical Linguistics.* Oxford: Blackwell.

Johnson, Daniel E. (2014) *Progress in Regression: Why Natural Language Data Calls For Mixed-Effects Models.* Manuscript, Lancaster University.

Johnson, Judith A. (1975) *A Transformational Analysis of the Syntax of Ælfric's Lives of Saints.* The Hague: Mouton.

Johnson, Kyle (2001) 'What VP Ellipsis can Do, and What it can't, but not Why.' In: Baltin, Marc and Chris Collins (eds.) *The Handbook of Contemporary Syntactic Theory.* Oxford: Blackwell, pp. 439–479.

Kastovsky, Dieter (2006) 'Vocabulary.' In: Hogg, Richard M. and David Denison (eds.) *A History of the English Language.* Cambridge: Cambridge University Press, pp. 199–311.

Kauhanen, Henri and George Walkden (2016) *Deriving the Constant Rate Effect.* Manuscript, Manchester University.

Kayne, Richard S. (1994) *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press.

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: Kemenade and Vincent (1997), pp. 326–352.

Kemenade, Ans van (1999) 'Sentential Negation and Clause Structure in Old English.' In: Tieken-Boon van Ostade, Ingrid, Gunnel Tottie and Wim van der Wurff (eds.) *Negation in the History of English.* Berlin: de Gruyter, pp. 147–165.

Kemenade, Ans van (2009) 'Discourse Relations and Word Order Change.' In: Hinterhölzl, Roland and Svetlana Petrova (eds.) *Information Structure and Language Change: New Approaches to Word Order Variation in Germanic.* Berlin: de Gruyter, pp. 91–118.

Kemenade, Ans van (2011) 'Secondary negation and information structure organization in the history of english' In: Larrivée, Pierre and Richard Ingham (eds.) *The Evolution of Negation: Beyond the Jespersen Cycle.* Berlin: de Gruyter, pp. 77–114.

Kemenade, Ans van and Bettelou Los (2006*a*) 'Discourse Adverbs and Clausal Syntax in Old and Middle English.' in *The Handbook of the History of English.* Kemenade and Los (2006*b*), pp. 224–248.

Kemenade, Ans van and Bettelou Los (eds.) (2006*b*) *The Handbook of the History of English.* Oxford: Wiley-Blackwell.

Kemenade, Ans van and Nigel Vincent (eds.) (1997) *Parameters of Morpho-Syntactic Change.* Cambridge: Cambridge University Press.

Kemenade, Ans van, Tanja Milicev and Harald Rolf Baayen (2008) 'The Balance between Syntax and Discourse in Old English.' In: Gotti, Maurizio, Marina Dossena and Richard Dury (eds.) *English Historical Linguistics 2006, Volume I: Syntax and Morphology.* Amsterdam: John Benjamins, pp. 3–22.

Kemenade, Ans van and Tanjaf Milicev (2011) 'Syntax and Discourse in Old and Middle English Word Order.' In: Jonas, Dianne, John Whitman and Andrew Garrett (eds.) *Grammatical Change: Origins, Nature, Outcomes.* Oxford: Oxford University Press, pp. 239–255.

Kim, Jong-Bok (2001) 'Negation, VP Ellipsis, and VP Fronting in English: A Construction-HPSG Analysis.' In: T'sou, Benjamin K., Olivia O. Kwong and Tom Lai (eds.) *Language, Information and Computation: Proceedings of The 15th Pacific Asia Conference 1–3 February 2001, Hong Kong.* HongKong: Hong Kong Language Information Sciences Research Centre, pp. 271–282.

Kim, Jong-Bok and Ivan A. Sag (2002) 'Negation without Head Movement.' *Natural Language and Linguistic Theory* **20.2**, 339–412.

Kivimaa, Kirsti (1966) *þe and þat as Clause Connectives in Early Middle English with Special Consideration of the Emergence of the Pleonastic þat.* Commentationes humanarum litterarum 39. Helsinki: Societas Scientiarum Fennic.

Kohonen, Viljo (1978) *On the Development of English Word Order in Religious Prose around 1000 and 1200 A.D.: A Quantitative Study of Word Order in Context.* Abo: Abo Akademi.

Koopman, Willelm F. (1995) 'Verb-Final Main Clauses in Old English Prose.' *Studia Neophilologica* **69**, 129–144.

Kroch, Anthony (1989) 'Reflexes of Grammar in Patterns of Language Change.' *Journal of Language Variation and Change* **1.3**, 199–244.

Kroch, Anthony (1994) 'Morphosyntactic Variation.' In: Beals, Kathryn (ed.) *Papers from the 30th Regional Meeting of the Chicago Linguistics Society.* Chicago: Chicago Linguistic Society, pp. 180–201.

Kroch, Anthony and Ann Taylor (1997) 'Verb Movement in Old and Middle English: Dialect Variation and Language Contact.' In: Kemenade and Vincent (1997), pp. 45–68.

Kroch, Anthony and Ann Taylor (2000) *Penn-Helsinki Parsed Corpus of Middle English. 2nd Edition.* `http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-3` (Accessed 10 April 2013) Department of Linguistics, University of Pennsylvania.

Kroch, Anthony, Ann Taylor and Ariel Diertani (2004) *Penn-Helsinki Parsed Corpus of Early Modern English.* `http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-2/index.html` (Accessed 20 April 2014) Department of Linguistics, University of Pennsylvania.

Kroch, Anthony, Beatrice Santorini and Ariel Diertani (2010) *Penn-Helsinki Parsed Corpus of Modern British English.* `http://www.ling.upenn.edu/hist-corpora/PPCMBE-RELEASE-1/index.html` (Accessed 30 May 2014) Department of Linguistics, University of Pennsylvania.

Kubota, Ai (2015) 'Transforming Manner Adverbs into Subject-Oriented Adverbs: Evidence from Japanese.' *Natural Language and Linguistic Theory* **33.3**, 1019–1046.

Kytö, Merja (1996) *Manual to the Diachronic Part of the Helsinki Corpus of English Texts: Coding Conventions and Lists of Source Texts.* Helsinki: University of Helsinki.

Labov, William (1994) *Principles of Linguistic Change. Volume 1: Internal Factors.* Oxford: Wiley-Blackwell.

Labov, William (2001) *Principles of Linguistic Change. Volume 2: Social Factors.* Oxford: Wiley-Blackwell.

Labov, William (2007) 'Transmission and Diffusion.' *Language* **83.2**, 344–387.

Labov, William (2010) *Principles of Linguistic Change. Volume 3: Cognitive and Cultural Factors.* Oxford: Wiley-Blackwell.

Lakoff, George (1991) 'Cognitive versus Generative Linguistics: How Commitments Influence Results.' *Language and Communication* **11.1/2**, 53–62.

Lenker, Ursula (2010) *Argument and Rhetoric: Adverbial Connectors in the History of English.* Topics in English Linguistics 64. Berlin: de Gruyter.

Levshina, Natalia (2015) *How to do Linguistics with R: Data Exploration and Statistical Analysis.* Amsterdam: John Benjamins.

Lewis, Robert E., Sherman M. Kuhn, Hans Kurath, John Reidy, David A. Jost, Helen W. Kao, Stephen F. Lappert, Lister M. Matheson, Ann Shannon and Mary Jane Williams (eds.) (1954–2001) *Middle English Dictionary.* Ann Arbor: The University of Michigan Press.

Light, Caitlin (2012) *The Syntax and Pragmatics of Fronting in Germanic.* PhD Dissertation, University of Pennsylvania.

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

Links, Meta and Ans van Kemande (2013) 'Correlative Constructions in Earlier English: The *þa* ... *þa* Construction.' *Paper delivered at SHES 11, Nijmegen, June 2013* .

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.

Los, Bettelou (2012) 'The Loss of Verb-Second and the Switch from Bounded to Unbounded Systems.' In: Meurman-Solin et al. (2012), pp. 21–46.

Los, Bettelou and Gea Dreschler (2012) 'The Loss of Local Anchoring: From Adverbial Local Anchors to Permissive Subjects.' In: Nevalainen, Tertu and Elizabeth C. Traugott (eds.) *The Oxford Handbook of the History of English.* Oxford: Oxford University Press, pp. 859–872.

Luzón Marco, María J. (1999) 'The Discursive Function of *do*-Support in Positive Clauses.' *Revista Española de Lingüística Aplicada* **13**, 87–103.

MacMullen, Ramsay (1988) *Corruption and the Decline of Rome.* New Haven, CT: Yale University Press.

Mahowald, Kyle (2011) *An LFG Approach to Old English Constituent Order.* MA Dissertation, University of Oxford.

Mair, Christian (2014) 'Do we Got a Difference? Divergent Developments of Semi-Auxiliary (have) got (to) in British and American English.' In: Hundt, Marianne (ed.) *Late Modern English Syntax.* Cambridge: Cambridge University Press, pp. 56–76.

Maling, Joan M. (1976) 'Notes on Quantifier-Postposing.' *Linguistic Inquiry* **7.4**, 708–718.

Martín, Javier C. and Antonio M. García (2007) 'On the Use of *ond*-Clauses in the Old English Gospels.' *Studia Neophilologica* **79**, 119–132.

Martinet, André (1952) 'Function, Structure and Sound Change.' *Word* **8**, 1–32.

McConnell-Ginet, Sally (1982) 'Adverbs and Logical Form: A Linguistically Realistic Theory.' *Language* **58.1**, 144–184.

McNally, Louise (2011) 'Existential Sentences.' In: Maienborn, Claudia, Klaus von Heusinger and Paul Portner (eds.) *Semantics: An International Handbook of Natural Language Meaning.* Volume 2. Berlin: de Gruyter, pp. 1829–1848.

Merchant, Jason (2001) *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis.* Oxford: Oxford University Press.

Merchant, Jason (2008*a*) 'An Asymmetry in Voice Mismatches in VP-Ellipsis and Pseudogapping.' *Linguistic Inquiry* **39.1**, 169–179.

Merchant, Jason (2008*b*) 'Variable Island Repair under Ellipsis.' In: Johnson, Kyle (ed.) *Topics in Ellipsis.* Cambridge: Cambridge University Press, pp. 132–153.

Merchant, Jason (2016) *Ellipsis: A Survey of Analytical Approaches.* Manuscript, University of Chicago.

Meurman-Solin, Anneli, María J. López-Courso and Bettelou Los (eds.) (2012) *Information Structure and Syntactic Change in the History of English.* Oxford: Oxford University Press.

Mitchell, Bruce (1964) 'Syntax and Word-Order in the Peterborough Chronicle 1122-1154.' *Neuphilologische Mitteilungen* **65**, 113–144.

Mitchell, Bruce (1985) *Old English Syntax.* 2 Volumes. Oxford: Clarendon.

Moberly, George H. (1881) *Venerabilis Baedae: Historia Ecclesiastica Gentis Anglorum.* London: Oxford University Press.

Montgomery, Michael (1992) 'The Etymology of *y'all.*' In: Hall, Joan H., Nick Doane and Dick Ringler (eds.) *Old English and New: Studies in Language and Linguistics in Honor of Frederic G. Cassidy.* New York: Garland, pp. 356–369.

Morris, Richard (1874-1893) *Cursor Mundi (The Cursur o the World). A Northumbrian Poem of the XIVth Century in Four Versions.* EETS o.s. 57, 59, 62, 66, 58, 99 & 101. London: Trübner & Co.

Mufwene, Salikoko, S. (2001) *The Ecology of Language Evolution.* Cambridge: Cambridge University Press.

Müller, Stefan and Stephen M. Wechsler (2014) 'Lexical Approaches to Argument Structure.' *Theoretical Linguistics* **40.1**, 1–76.

Myler, Neil J. (2014) *Building and Interpreting Possession Sentences.* PhD Dissertation, New York, NY: UMI Dissertation Publishing.

Nielsen, Hans F. (2003) 'On the Demise of Old English.' In: Heizmann, Wilhelm and Astrid van Nahl (eds.) *Runica - Germanica - Mediaevalia.* Berlin: de Gruyter, pp. 496–508.

Nielsen, Leif A. (2004) 'Verb Phrase Ellipsis Detection Using Automatically Parsed Text.' *Proceedings of the 20th International Conference on Computational Linguistics*, pp. 1093–1099.

Ohkado, Masayuki (2004) 'Coordinate Clauses in Old English with Special Reference to Ælfric's Catholic Homilies.' *Folia Linguistica Historica* **251.2**, 155–176.

Osborne, Timothy (2013) 'The Distribution of Floating Quantifiers: A Dependency Grammar Analysis.' *Proceedings of the International Conference on Dependency Linguistics* **2**.

Otoguro, Ryu (2006) 'Dissociation between Phrase-Structure and Functions: A Lexical-Functional Analysis of English Negator *not.*' *Osaka University Papers in English Linguistics* **11**, 85–109.

Park, Ji Seon (2008) *Negative Yes/No Question-Answer Sequences in Conversation: Grammar, Action, and Sequence Organization.* PhD Dissertation, University of California, Los Angeles.

Peng, Rui (2012) 'Critical Frequency as an Independent Variable in Grammaticalization.' *Studies in Language* **36.2**, 345–381.

Perlmutter, David M. (1971) *Deep and Surface Constraints in Syntax.* New York: Holt, Rinehart and Winston.

Perlmutter, David M. and Paul M. Postal (1983) 'Some Proposed Laws of Basic Clause Structure.' In: Perlmutter, David M. (ed.) *Studies in Relational Grammar 1.* Chicago: University of Chicago Press, pp. 81–128.

Pintzuk, Susan (1993) 'Verb Seconding in Old English: Verb Movement to Infl.' *The Linguistic Review* **10**, 5–35.

Pintzuk, Susan (1995) 'Variation and Change in Old English Clause Structure.' *Language Variation and Change* **7**, 229–260.

Pintzuk, Susan (1996) 'Cliticization in Old English.' In: Halpern, Aaron L. and Arnold Zwicky (eds.) *Approaching Second: Second Position Clitics and Related Phenomena.* Stanford, CA: CSLI Press, pp. 366–400.

Pintzuk, Susan (1999) *Phrase Structures in Competition: Variation and Change in Old English Word Order.* New York: Garland.

Pintzuk, Susan (2005) 'Arguments against a Universal Base: Evidence from Old English.' *English Language and Linguistics* **9**, 115–138.

Pintzuk, Susan and Eric Haeberli (2008) 'Structural Variation in Old English Root Clauses.' *Language Variation and Change* **20**, 367–407.

Pintzuk, Susan, George Tsoulas and Anthony Warner (eds.) (2000) *Diachronic Syntax, Models and Mechanisms: Proceedings of DiGS 5.* Oxford: Oxford University Press.

Poletto, Cecilia (2014) *Word Order in Old Italian.* Oxford: Oxford University Press.

Pollock, Jean-Yves (1989) 'Verb Movement, Universal Grammar, and the Structure of IP.' *Linguistic Inquiry* **20.3**, 365–424.

Popper, Karl R. (1963) *Conjectures and Refutations.* London: Routledge & Kegan Paul.

Potsdam, Eric (1997) 'English Verbal Morphology and VP Ellipsis.' *Proceedings of NELS* **27**, 354–368.

Potsdam, Eric (1998) 'A Syntax for Adverbs.' In: Gelderen, Elly van and Vida Samiin (eds.) *The Proceedings of the Twenty-Seventh Western Conference on Linguistics.* Fresno, CA: California State University Press, pp. 397–411.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartik (1972) *A Grammar of Contemporary English.* London: Longman.

Radford, Andrew (2004) *English Syntax: An Introduction.* Cambridge: Cambridge University Press.

Ramchand, Gillian and Peter Svenonius (2014) 'Deriving the Functional Hierarchy.' *Language Sciences* **46.2**, 152–174.

Randall, Beth (2004) *CorpusSearch 2.* `http://corpussearch.sourceforge.net/` (Accessed 7 November 2013) Sourceforge.net.

Repp, Sophie (2009) *Negation in Gapping.* Oxford: Oxford University Press.

Rhee, Seongha (2002) 'Semantic Change of English Preposition *against*: A Grammaticalization Perspective.' *Language Research* **38.2**, 563–583.

Rickford, John R. (2001) *African American Vernacular English.* Oxford: Blackwell.

Ringe, Don (2006) *From Proto-Indo-European to Proto-Germanic.* Oxford: Oxford University Press.

Ringe, Don and Ann Taylor (2014) *The Development of Old English.* Volume 2. Oxford: Oxford University Press.

Rizzi, Luigi (1996) 'Residual Verb Second and the *Wh*-Criterion.' In: Rizzi, Luigi and Adriana Belletti (eds.) *Parameters and Functional Heads.* Oxford: Oxford University Press, pp. 63–90.

Rizzi, Luigi (1997) 'The Fine Structure of the Left Periphery.' In: Haegeman, Liliane (ed.) *Elements of Grammar.* Dordrecht: Kluwer, pp. 281–337.

Roberts, Ian G. (1985) 'Agreement Parameters and the Development of English Modal Auxiliaries.' *Natural Language and Linguistic Theory* **3**, 21–58.

Roberts, Ian G. (2010) *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals.* Cambridge, MA: MIT Press.

Roberts, Ian G. and Anna Roussou (2003) *A Minimalist Approach to Grammaticalization.* Cambridge: Cambridge University Press.

Robson, Karen and David Pevalin (2016) *Multilevel Modeling in Plain Language.* Los Angeles: Sage.

Rochette, Anne (1990) 'The Selectional Properties of Adverbs.' In: Ziolkowski, Michael, Manuela Noske and Karen Deaton (eds.) *Papers from the 26th Regional Meeting of the Chicago Linguistic Society, Volume 1: The Main Session.* Chicago: Chicago Linguistic Society, pp. 379–391.

Roeper, Thomas (1999) 'Universal Bilingualism.' *Bilingualism: Language and Cognition* **2.3**, 169–186.

Rohrbacher, Bernhard W. (1999) *Morphology-Driven Syntax: A Theory of V to I Raising and Pro-Drop.* Los Angeles: Sage.

Roth, Wilhelm (1914) *Die Wortstellung im Aussage-Hauptsatz Angelsächsischer Originalliteratur.* PhD Dissertation, Friedrich-Wilhelms-Universität Berlin.

Ruys, Eddy G. and Yoad Winter (2011) 'Quantifier Scope in Formal Linguistics.' In: Gabbay, Dov M. and Franz Guenthner (eds.) *Handbook of Philosophical Logic.* Dordrecht: Kluwer Academic Publishers, pp. 159–225.

Sadler, Louisa (1998) 'English Auxiliaries as Tense Inflections.' *Essex Research Reports* **24**, 1–16.

Sadler, Louisa (2005) 'Function Spreading in Coordinate Structures.' *Lingua* **116.1**, 1777–1806.

Sag, Ivan A. (1976) *Deletion and Logical Form.* PhD Dissertation, Massachusetts Institute of Technology.

Sag, Ivan A. (1978) 'Floating quantifiers, Adverbs and Extraction Sites.' *Linguistic Inquiry* **9**, 146–150.

Saitz, Robert L. (1955) *Functional Word Order in Old English Subject-Object Patterns.* PhD Dissertation, University of Wiscounsin.

Saussure, Ferdinand de. (1916/2014) *Cours de Linguistique Générale: Studienausgabe in deutscher Sprache.* Tübingen: Narr.

Scalise, Sergio and Emiliano Guevara (2005) 'The Lexicalist Approach to Word-Formation and the Notion of the Lexicon.' In: Štekauer and Lieber (2005), pp. 147–187.

Sells, Peter (2005) 'The Peripherality of the Icelandic Expletive.' *Proceedings of the LFG Conference* **10**, 408–428.

Shieber, Stuart M., Mary Dalrymple and Fernando C. N. Pereira (1996) 'Interactions of Scope and Ellipsis.' *Linguistics and Philosophy* **19.5**, 527–552.

Siddiqi, Daniel (2014) 'Word-Formation and Syntax.' In: Carnie, Andrew, Daniel Siddiqi and Yosuke Sato (eds.) *The Routledge Handbook of Syntax.* Oxford: Routledge, pp. 345–364.

Smith, Alphonso C. (1893) 'The Order of Words in Anglo-Saxon Prose.' *Publications of the Modern Language Association of America* **8**, 210–244.

Spector, Ilona (2009) 'Hebrew Floating Quantifiers.' *Proceedings of the LFG Conference* **14**, 520–540.

Spencer, Andrew (2005) 'Word-Formation and Syntax.' In: Štekauer and Lieber (2005), pp. 73–97.

Speyer, Augustin (2010) *Topicalization and Stress Clash Avoidance in the History of English.* Berlin: de Gruyter.

Sportiche, Dominique (1988) 'A Theory of Floating Quantifiers and its Corollaries for Constituent Structure.' *Linguistic Inquiry* **19**, 425–449.

Štekauer, Pavel and Rochelle Lieber (eds.) (2005) *Handbook of Word-Formation.* Amsterdam: Springer.

Stockwell, Robert P. and Donka Minkova (1990) 'Verb Phrase Conjunction in Old English.' In: Andersen, Henning and Konrad Koerner (eds.) *Historical Linguistics 1987: Papers from the 8th International Conference on Historical Linguistics.* Amsterdam: John Benjamins, pp. 499–515.

Tagliamonte, Sali A., Alexandra D'Arcy and Bridget Jankowski (2010) 'Social Work and Linguistic Systems: Marking Possession in Canadian English.' *Language Variation and Change* **22**, 149–173.

Taylor, Ann, Anthony Warner, Susan Pintzuk and Frank Beths (2003) *The York-Toronto-Helsinki Corpus of Old English Prose (YCOE).* `http://www-users.york.ac.uk/ lang22/YCOE/YcoeHome.htm` (Accessed 10 April 2013) Oxford Text Archive.

Tenney, Frank (1916) 'Race Mixture in the Roman Empire.' *The American Historical Review* **21.4**, 689–708.

Trask, Robert L. (1996) *Historical Linguistics.* London: Arnold.

Traugott, Elizabeth C. (1972) *A History of English Syntax: A Transformational Approach to the History of English Sentence Structure.* New York: Holt, Rinehart and Winston.

Traugott, Elizabeth C. (1992) 'Old English Syntax.' In: Hogg, Richard M. (ed.) *The Cambridge History of the English Language.* Cambridge: Cambridge University Press, pp. 168–289.

Travis, Lisa (1988) 'The Syntax of Adverbs.' *McGill Working Papers in Linguistics: Special Issue on Comparative Germanic Syntax*, pp. 280–310.

Trips, Carola (2002) *From OV to VO in Early Middle English.* Amsterdam: John Benjamins.

Trudgill, Peter, Terttu Nevalainen and Ilse Wischer (2002) 'Dynamic *have* in North American and British Isles English.' *English Language and Linguistics* **6**, 1–15.

Varela Pérez, José Ramón (2007) 'Negation of Main Verb *have*: Evidence of a Change in Progress in Spoken and Written British English.' *Neuphilologische Mitteilungen* **108.1**, 223–246.

Vikner, Sten (1997) 'V-to-I Movement and Inflection for Person in All Tenses.' In: Haegeman, Liliane (ed.) *The New Comparative Syntax.* London: Longman, pp. 189–213.

Walkden, George (2014) *Syntactic Reconstruction and Proto-Germanic.* Oxford: Oxford University Press.

Wallenberg, Joel C. (2009) *Antisymmetry and the Conservation of C-Command: Scrambling and Phrase Structure in Synchronic and Diachronic Perspective.* PhD Dissertation, University of Pennsylvania.

Wallenberg, Joel C. (2014) *Extraposition is Disapppearing.* Manuscript, Newcastle University.

Ward-Perkins, Bryan (2006) *The Fall of Rome: And the End of Civilization.* Oxford: Oxford University Press.

Warner, Anthony R. (1985) *The Structure of English Auxiliaries: A Phrase Structure Grammar.* Bloomington: Indiana University Linguistics Club.

Warner, Anthony R. (2006) 'Variation and the Interpretation of Change in Periphrastic Do.' In: Kemenade and Los (2006*b*), pp. 45–67.

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.

Wårvik, Brita (1995) 'The Ambiguous Adverbial/Conjunctions *Tha* and *Thonne* in Middle English: A Discourse-Pragmatic Study of *Then* and *When* in Early English Saints' Lives.' In: Jucker, Andreas H. (ed.) *Historical Pragmatics.* Amsterdam: John Benjamins, pp. 345–357.

Wårvik, Brita (2011) 'Connective or 'Disconnective' Discourse Marker? Old English *þa*, Multifunctionality and Narrative Structuring.' In: Meurman-Solin, Anneli and Ursula Lenker (eds.) *Connectives in Synchrony and Diachrony in European Languages.* http://www.helsinki.fi/varieng/series/volumes/08/warvik/ (Accessed January 2017) Helsinki: University of Helsinki peer-reviewed online publications.

Wasow, Thomas (1997) 'Remarks on Grammatical Weight.' *Language Variation and Change* **9.1**, 81–105.

Wasow, Thomas and Jennifer Arnold (2003) 'Post-verbal Constituent Ordering in English.' In: Rohdenburg, Günther and Britta Mondorf (eds.) *Determinants of Grammatical Variation in English.* Topics in English Linguistics (TiEL) 43 Berlin: de Gruyter, pp. 119–154.

Weinreich, Uriel, William Labov and Marvin Herzog (1968) 'Empirical foundations for a theory of language change' In: Lehmann, Wilfried P. and Yakov Malkiel (eds.) *Directions for historical linguistics: a symposium.* Austin: University of Texas Press, pp. 95–198.

Wilder, Chris (1997*a*) 'English Finite Auxiliaries in Syntax and Phonology.' In: Black, James R. and Virginia Motapanyane (eds.) *Clitics, Pronouns and Movement.* Amsterdam: John Benjamins, pp. 321–362.

Wilder, Chris (2013) ‘English ‘Emphatic *do*’.’ *Lingua* **128**, 142–171.

Wilder, Chris (ed.) (1997*b*) *The Role of Economy Principles in Linguistic Theory.* Berlin: Akademie-Verlag.

Williams, Alexander (2000) ‘Null Subjects in Middle English Existentials.’ In: Pintzuk et al. (2000), pp. 164–187.

Winford, Donald (1996) ‘The Problem of Syntactic Variation.’ In: Arnold, Jennifer, Renée Blaker, Brad Davidson, Scott Schwenter and Julie Solomon (eds.) *Sociolinguistic Variation: Data, Theory and Analysis. Selected Papers from NWAVE 23.* Stanford: CSLI, pp. 177–192.

Wright, Thomas and James O. Halliwell-Phillipps (1845) *Reliquiae Antiquae.* London: Smith.

Wyner, Adam (1998) ‘A Discourse Theory of Manner and Factive Adverbial Modification.’ In: Hulstij, Joris and Anton Nijholt (eds.) *The Proceedings of Twendial ‘98: Formal Semantics and Pragmatics of Dialogue.* Enschede: University of Twente Press, pp. 249–267.

Yamakawa, Kikuo (1969) ‘The Development of *when* as Subordinate Conjunction or Relative Adverb.’ *Hitotsubashi Journal of Arts and Sciences* **10.1**, 8–42.

Yang, Charles (2002) *Knowledge and Learning in Natural Language.* Oxford: Oxford University Press.

Zeijlstra, Hedde (2015) ‘The Morpho-Syntactic Realisation of Negation.’ In: *Syntax - Theory and Analysis. International Handbook.* Volume 1. Berlin: de Gruyter, pp. 274–309.

Zimmermann, Richard (2009) *Topics and Pronouns in the Clausal Left Periphery in Old English.* MA Dissertation, University of York.

Zimmermann, Richard (2012) ‘*Self* as a Non-Postposing Element in Old English.’ *Generative Grammar in Geneva* **8**, 39–58.

Zimmermann, Richard (2014*a*) *Dating Hitherto Undated Old English Texts based on Text-Internal Criteria.* Manuscript, University of Geneva.

Zimmermann, Richard (2014*b*) ‘Distributional Differences between Old English Main Clauses with and without a Conjunction.’ *Proceedings of the LFG Conference* **19**, 566–585.

Zimmermann, Richard (2015) *The Parsed Corpus of Middle English Poetry.* `http://www.pcmep.net/` (Accessed May 2016) University of Geneva.

Zoëga, Geir T. (1910) *A Concise Dictionary of Old Icelandic.* Oxford: Clarendon.

Zyman, Erik (2012) *Two Investigations of Adverbs and Clause Structure in English.* MA Dissertation, University of Princeton.

# Appendices

# Appendix A

# Negative Inversion in Modern English

This appendix presents an explicit analysis of Modern English negative inversion in LFG. An adequate treatment of this topic requires paying some attention to its semantics. The model of the syntax-semantics interface assumed here is Glue Semantics, which has become the standard for work on semantics within LFG (for introductions to Glue Semantics, see Dalrymple 2001, Asudeh 2012). The meaning language used is simple first order logic extended with lambda calculus.

Certain Modern English lexical items are assumed to decompose into a negation feature, `POL=neg`, and some other, simpler operator. The semantic contribution of the relevant lexical items can be paraphrased as in (1).

(1)
a. never = not ever
b. nor = not or
c. neither = not either
d. no = not some
e. nowhere = not somewhere
f. nobody = not somebody
g. nothing = not something

As explained in the main text of chapter 2, the isolated negation feature can be targeted by syntactic rules that license negative constituents in the specifier of $\texttt{CP}_{MAT}$ and trigger subject-auxiliary inversion via the introduction of a non-canonical declarative clause type feature. The sketches of the relevant rules are repeated in (2a) and (2b).

(2)
a. $$\texttt{CP}_{MAT} \rightarrow \begin{array}{c} \texttt{XP} \\ \texttt{...} \\ (\downarrow \texttt{POL}) =_c \texttt{neg} \\ (\uparrow \texttt{CLAUSE-TYPE}) = \texttt{non-can-decl} \\ \texttt{...} \end{array} \quad \begin{array}{c} \texttt{C'}_{MAT} \\ \uparrow = \downarrow \end{array}$$

b. $$\texttt{C'}_{MAT} \rightarrow \begin{array}{c} \texttt{C}_{MAT} \\ \uparrow = \downarrow \end{array} \quad \begin{array}{c} \texttt{IP} \\ \uparrow = \downarrow \\ (\uparrow \texttt{CLAUSE-TYPE}) =_c \texttt{non-can-decl} \end{array}$$

The negative polarity feature could be regarded either as a functionally motivated mnemonic label for some property, such as non-positivity or non-predicate focus, which subject-auxiliary inversion sentences (tend to) share (Goldberg 2006, 2009), or as a purely formalistic, “autonomous” feature used to engineer some practicable inversion mechanism (e.g., Borsley and Newmeyer 2009). The analysis does not hinge on this detail.

In terms of semantics, the negative operator can either be assigned a meaning on its own, or the isolated negation feature and the rest-operator can be given two distinct semantic denotations. In the former case, the denotations of the negative lexical items can belong to various semantic types. In the latter case, the isolated negation feature will be assumed to scope over the propositional mother of the function in which it is contained. It will invert the truth values of that proposition, i.e., belong to type $< t, t >$.

I will now illustrate how exactly the analysis sketched above works in specific instances of negative inversion. First, I will discuss an example that involves the clause-initial negative adverb *never*. As explained in the general outline above, the meaning of *never* will be divided into a negation and an operator part. The operator is represented simply by its `PRED` feature, whose mnemonic label is reversed for convenience from "never" to "at all times." The negator is represented by a separate negation feature, `POL=neg`. The disentangled lexicon entry for *never* is presented in (3a). A sketch of *never* in an f-structure configuration is shown in (3b), where, as would typically be the case, it functions as an adjunct.

(3) a.

$$\textit{never}\ \texttt{ADV}\quad \begin{array}{l} (\uparrow \texttt{PRED}) = \texttt{'AT-ALL-TIMES'} \\ (\uparrow \texttt{POL}) = \texttt{neg} \end{array}$$

b.

$$\left[\begin{array}{ll} \ldots & \\ \textsc{subj} & [\quad] \\ \textsc{adjunct} & \left\{ \left[\begin{array}{ll} \textsc{pred} & \text{'AT-ALL-TIMES'} \\ \textsc{pol} & \text{neg} \end{array}\right] \right\} \\ \ldots & \end{array}\right]$$

Following the general format for rules parsing negative initial constituents in (2a), I propose the following specific phrase structure rule to handle initial negative adverbs such as *never*, (4).

(4)

$$\texttt{CP}_{MAT} \rightarrow \quad \begin{array}{c} \texttt{ADVP} \\ (\uparrow \texttt{UDF}) = \downarrow \\ (\uparrow \texttt{UDF}) \in (\uparrow \texttt{ADJUNCT}) \\ (\downarrow \texttt{POL}) =_{c} \texttt{neg} \\ (\uparrow \texttt{CLAUSE-TYPE}) = \texttt{non-can-decl} \end{array} \qquad \begin{array}{c} \texttt{C'}_{MAT} \\ \uparrow = \downarrow \\ (\downarrow \texttt{TENSE}) \\ \\ \\ \end{array}$$

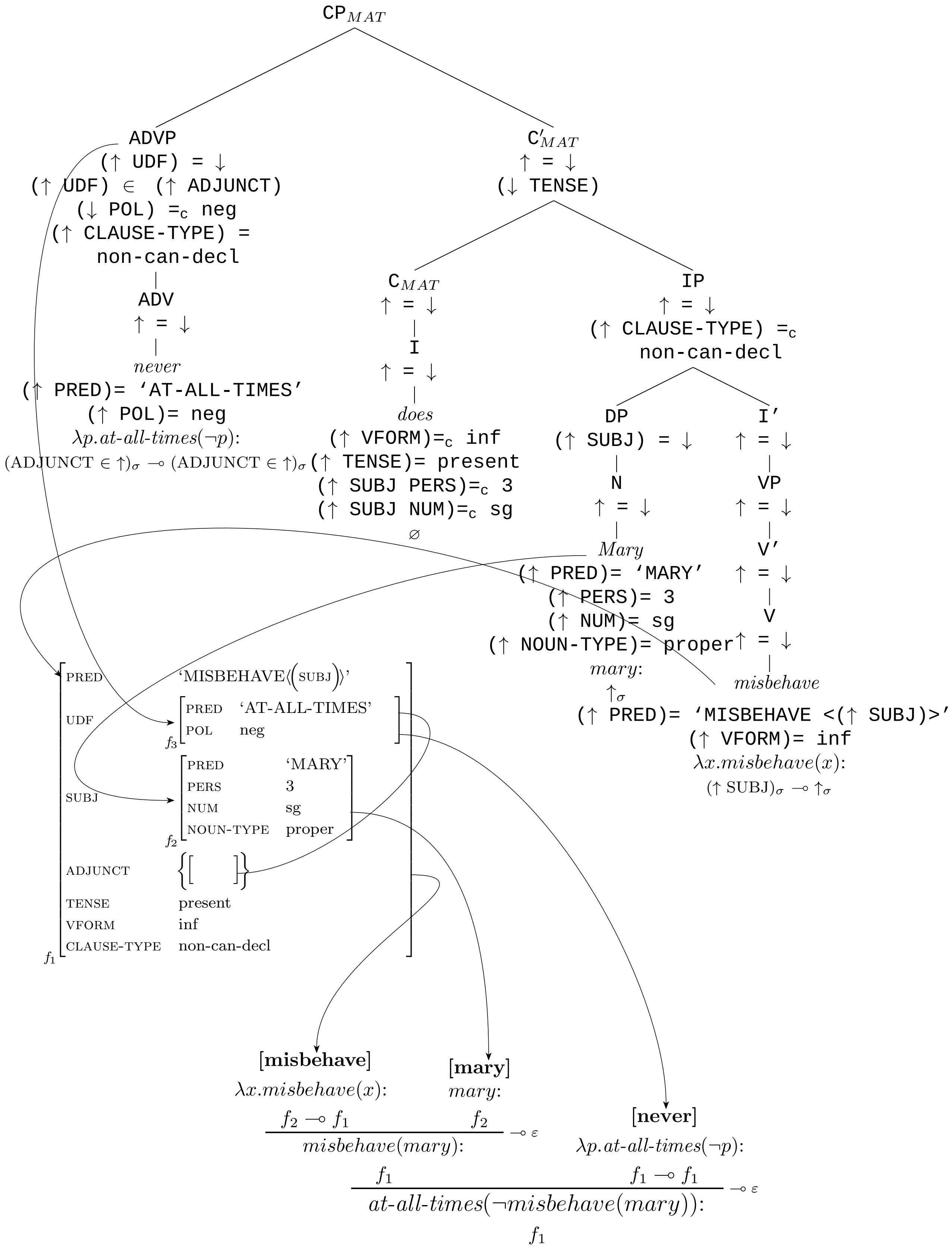
An adverb phrase, `ADVP`, with a negation feature can be parsed as the specifier of $\texttt{CP}_{MAT}$, where it will be mapped onto a discourse function, `UDF`, perhaps some kind of focus, and unify with a member of the local adjunct set. The application of this rule introduces a non-canonical declarative clause type feature into the structure.

Although the item *never* consists of two distinct syntactic features, its meaning remains unified. Its semantic denotation simply states that at all times, the proposition it modifies does not hold. Selection of the correct proposition is ensured by meaning constructors. Here, a glue derivation will consume the semantic resource associated with the f-structure mother of the adjunct function of which *never* is an element to produce a new, modified meaning that is associated with the same kind of semantic resource. The semantic type of *never* is thus $< t, t >$. The semantic contribution of *never* is shown in (5). The denotation is followed by a colon and the meaning constructor.

(5) $[\textbf{never}]\ \lambda p.\textit{at-all-times}(\neg p)\text{:}\quad (\text{ADJUNCT} \in \uparrow)_{\sigma} \multimap (\text{ADJUNCT} \in \uparrow)_{\sigma}$

It is now possible to parse sentences with subject-auxiliary inversion after the negative, clause-initial adverb *never*. As an example, I will discuss the sentence *Never does Mary misbehave*. For reasons of clarity, a maximally explicit parse of this sentence will be presented. It involves the following aspects: (i) the full lexicon entries for every word, including their syntactic features, semantic denotations and meaning constructors, (ii) the integration of the lexical items into, and the construction of, a hierarchical phrase structure tree, as well as its phi-projection to a regulatory array of features, (iii) the sigma-projection from the feature structure to the semantic contributions of every f-structure node, and their involvement in a proof-theoretic deduction of a formula representing the meaning of the whole sentence. In short, the analysis involves the mappings lexicon $\rightarrow$ c-structure, c-structure $\rightarrow$ f-structure, and f-structure $\rightarrow$ semantics. The complete parse is shown in (6).

(6) Never does Mary misbehave.



The initial constituent *never* contains a syntactic negation feature, POL=neg. It can thus only be parsed by the CP-rule in (4), whose application will simultaneously introduce a non-canonical declarative clause type feature. This feature, in turn, is expected by the C′-rule in (61), which thus triggers subject-auxiliary inversion. The semantic composition of the sentence is straightforward, as shown by the proof tree. The general meaning constructors of the lexical items are instantiated, mediated through the syntactic parse, to specific f-structure nodes. For example, *Mary*'s meaning constructor $\uparrow_\sigma$ is instantiated to the f-structure that is headed by its PRED feature, i.e., here to the subject function, $f_2$. Similarly, the meaning constructor of the modifier *never*, $(\text{ADJUNCT} \in \uparrow)_\sigma \multimap (\text{ADJUNCT} \in \uparrow)_\sigma$ , gets instantiated, based on the particular syntactic parse shown, to $f_1 \multimap f_1$. The item *does* is assumed to be semantically vacuous. The model produces one unique solution for the syntactic parse of the sentence and can compose exactly one semantic formula representing its meaning.

The examples I will discuss next involve clause-initial PPs whose complements are quantified by the negative quantifier *no*. I will first present the syntactic features in the lexicon entry for *no*, then define phrase structure rules for the construction of PPs and for their integration into the clausal structure, and finally list the semantic contribution of the word *no* and the relevant combinatorial rules.

First the syntactic features in *no*'s lexicon entry are the following. As sketched in (1), the negative operator *no* is split into a quantifier, whose PRED-label I invert for convenience from "no" to "some," (↑ PRED)='SOME', and a syntactic negation feature, POL=neg. I assume that quantifiers head their own function, quantification, which I label QUANT (sometimes called SPEC elsewhere). The isolated negation feature is introduced somewhere within the f-structure that contains the quantification function. Within prepositional phrases two distinct placements of the negation feature are possible. Negation can scope over the f-structure that immediately contains quantification, hence the nominal complement itself. This is achieved by means of the inside-out functional uncertainty constraint ((QUANT ↑) POL)= neg. Alternatively, the negative feature can apply to the function immediately dominating the nominal complement, hence the f-structure referring to the domain of the whole PP. Complements of prepositions function as semantically restricted objects, $\text{OBJ}_{th}$. The desired, alternative feature placement can therefore be achieved by the inside-out functional uncertainty constraint (($\text{OBJ}_{th}$ QUANT ↑) POL)= neg. The proper interpretation of the negation feature requires that it be applied to the semantic contribution of an f-structure that contains a subject, hence a proposition. This requirement might follow directly from a richer meaning language (e.g., events). With the simple logical representation of meaning used here, however, a syntactic constraint must be used to force the presence of a subject in the right place. Specifically, an inside-out functional uncertainty constraint will ensure that the f-structure containing any f-structure with the negation feature, abbreviated here as ANY, also contains a subject, SUBJ. The label ANY should be understood as a meta-category subsuming all those grammatical functions that can be realized as prepositional phrases, like XADJUNCT, ADJUNCT etc. The resultant lexicon entry for *no* is presented in (7).

(7) *no* Q (↑ PRED)= 'SOME'

{ ((QUANT ↑) POL)= neg
((ANY QUANT ↑) SUBJ)
| (($\text{OBJ}_{th}$ QUANT ↑) POL)= neg
((ANY $\text{OBJ}_{th}$ QUANT ↑) SUBJ) }

To clarify the workings of these lexical constraints, I show in (8) below two skeletal f-structures for PPs whose complements are negated by the quantifier *no*. The negation feature is placed in two different positions according to the two options for negation placement explained above. The f-structure in (8a) is derived from the first disjunct of the set of constraints for the introduction of the negation feature. It places the feature within the f-structure that embeds the quantification function, QUANT, hence the complement of the preposition, $\text{OBJ}_{th}$. It is licensed there because the function of the PP, here a member of an XADJUNCT set, labeled $a_1$, also includes a subject function, SUBJ, labeled $a_2$. The f-structure in (8b) makes use of the second disjunct of the set of constraints for the introduction of the negation feature. In this structure, the negation feature is placed within the f-structure that contains the complement of the preposition, $\text{OBJ}_{th}$. In other words, the negation feature will be found in the outer f-structure corresponding to the whole PP. The feature is licensed in this position because the matrix f-structure that contains the function corresponding to the PP, labeled $b_1$, also includes a subject, SUBJ, labeled $b_2$.

(8) a.

$$\begin{bmatrix} \text{PRED} & \text{'...'} \\ \text{SUBJ} & [\;\;] \\ \text{XADJUNCT} & \left\{ {}_{a_1}\begin{bmatrix} \text{PRED} & \text{'PREPOSITION}\langle...\rangle\text{'} \\ \text{SUBJ} & {}_{a_2}[\;\;] \\ \text{OBJ}_{th} & \begin{bmatrix} \text{PRED} & \text{'NOUN'} \\ \text{QUANT} & \begin{bmatrix}\text{PRED} & \text{'SOME'}\end{bmatrix} \\ \text{POL} & \text{neg}\end{bmatrix}\end{bmatrix}\right\} \\ \ldots & \end{bmatrix}$$

b.

$${}_{b_1}\begin{bmatrix} \text{PRED} & \text{'...'} \\ \text{SUBJ} & {}_{b_2}[\;\;] \\ \text{XADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{'PREPOSITION}\langle...\rangle\text{'} \\ \text{SUBJ} & [\;\;] \\ \text{OBJ}_{th} & \begin{bmatrix} \text{PRED} & \text{'NOUN'} \\ \text{QUANT} & \begin{bmatrix}\text{PRED} & \text{'SOME'}\end{bmatrix}\end{bmatrix} \\ \text{POL} & \text{neg}\end{bmatrix}\right\} \\ \ldots & \end{bmatrix}$$

Next, I will outline the syntactic rules necessary to parse clause-initial negatively quantified constituents. The phrase structure rules in (9) allow the construction of `PP`s with quantified complements.

(9)

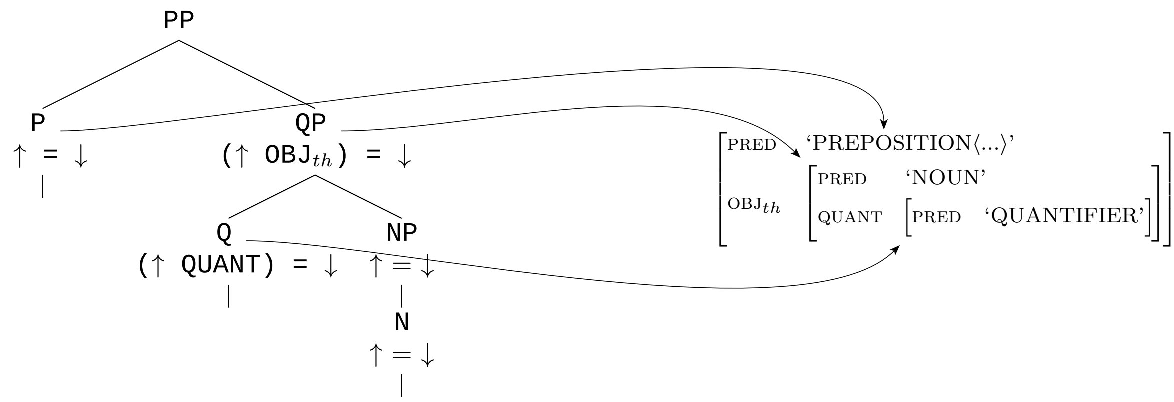
$$\text{a. } \texttt{PP} \rightarrow \underset{\uparrow\,=\,\downarrow}{\texttt{P}} \quad \underset{(\uparrow\ \texttt{OBJ}_{th})\,=\,\downarrow}{\texttt{QP}}$$

$$\text{b. } \texttt{QP} \rightarrow \underset{(\uparrow\ \texttt{QUANT})\,=\,\downarrow}{\texttt{Q}} \quad \underset{\uparrow\,=\,\downarrow}{\texttt{NP}}$$

$$\text{c. } \texttt{NP} \rightarrow \underset{\uparrow\,=\,\downarrow}{\texttt{N}}$$

The complement of a preposition functions as a thematically restricted object, $\texttt{OBJ}_{th}$. The quantifier heads a quantification function, `QUANT`. A skeleton phrase structure tree for a `PP` constructed from the rules in (9) is presented in (10a), and the corresponding functional structure in (10b).

(10) a. b.




A quantified nominal would be constructed in a similar way except that the outer `PP`-layer would be absent and some other phrase structure rule would determine its grammatical function.

I will now turn to the phrase structure rules necessary to parse negative clause-initial `PP`s. If the negation feature introduced by *no* is interpreted within the f-structure referring to the complement of the preposition, as illustrated in (8a), the phrase must be parsed with the rule in (11). The rule adjoins the `PP` to $\texttt{CP}_{MAT}$. Further, the rule does not permit the placement of a negation feature on the outer f-structure referring to the domain of the whole `PP`, as in (8b), via the constraint $(\downarrow\ \texttt{POL}) \neq \texttt{neg}$. The application of the rule maps the initial `PP` onto an arbitrarily deeply embedded open adjunct function with an externally controlled subject, $\{\texttt{XCOMP|COMP}\}^{*}\ \texttt{XADJUNCT}$. Another grammatical function for the `PP`, like `ADJUNCT`, is not viable because the constraints in the lexicon entry of *no* require the presence of a subject within the function containing the negated function, hence an open grammatical function.

(11)

$$\texttt{CP}_{MAT} \rightarrow \underset{\substack{\downarrow\,\in\,(\uparrow\ \{\texttt{XCOMP|COMP}\}^{*}\ \texttt{XADJUNCT}) \\ (\uparrow\ \texttt{XADJUNCT SUBJ}) = (\uparrow\ \texttt{SUBJ}) \\ (\downarrow\ \texttt{OBJ}_{th}\ \texttt{PRON-TYPE}) \neq \texttt{wh} \\ (\downarrow\ \texttt{POL}) \neq \texttt{neg}}}{\texttt{PP}} \quad \underset{\uparrow\,=\,\downarrow}{\texttt{CP}_{MAT}}$$

If, on the other hand, the negation feature percolates up to the level of the whole PP, as illustrated in (8b), the phrase must be parsed with the rules in (12) or (13). These rules are based on the general format in (60). Hence, these rules place the PP in the specifier of $\mathtt{CP}_{MAT}$ and introduce a non-canonical declarative clause type feature, which then triggers the subject-auxiliary inversion mechanism. The two rules preclude phrases in their specifier without a negation feature on their corresponding outer f-structure, as for example in (8a), since the presence of such a feature is required by the constraint $(\downarrow \mathtt{POL}) =_c \mathtt{neg}$. Both rules interpret the initial PP as an underspecified discourse function, UDF. The former unifies the UDF with a member of the local XADJUNCT set, whose subject is externally controlled, the latter with a member of the local ADJUNCT set, which does not include a subject function. Either grammatical function is possible here because the negation feature now refers to the whole PP, and not to the prepositional complement, which renders moot the constraints in the lexicon entry of *no* regarding the obligatory occurrence of a subject within the range of the PP. The grammatical function of the f-structure corresponding to the whole PP may thus be open or closed.

(12) $\mathtt{CP}_{MAT} \rightarrow \begin{array}{c} \mathtt{PP} \\ (\uparrow \mathtt{UDF}) = \downarrow \\ (\uparrow \mathtt{UDF}) \in (\uparrow \mathtt{XADJUNCT}) \\ (\uparrow \mathtt{XADJUNCT\ SUBJ}) = (\uparrow \mathtt{SUBJ}) \\ (\downarrow \mathtt{POL}) =_c \mathtt{neg} \\ (\uparrow \mathtt{CLAUSE\text{-}TYPE}) = \mathtt{non\text{-}can\text{-}decl} \end{array} \quad \begin{array}{c} \mathtt{C'}_{MAT} \\ \uparrow = \downarrow \\ (\downarrow \mathtt{TENSE}) \\ \\ \\ \\ \end{array}$

(13) $\mathtt{CP}_{MAT} \rightarrow \begin{array}{c} \mathtt{PP} \\ (\uparrow \mathtt{UDF}) = \downarrow \\ (\uparrow \mathtt{UDF}) \in (\uparrow \mathtt{ADJUNCT}) \\ (\downarrow \mathtt{POL}) =_c \mathtt{neg} \\ (\uparrow \mathtt{CLAUSE\text{-}TYPE}) = \mathtt{non\text{-}can\text{-}decl} \end{array} \quad \begin{array}{c} \mathtt{C'}_{MAT} \\ \uparrow = \downarrow \\ (\downarrow \mathtt{TENSE}) \\ \\ \\ \end{array}$

Lastly, I will discuss the semantics of the quantifier *no* and the semantic contribution of some relevant phrase structure rules. Unlike the semantics of *never*, the semantics of *no* mirrors the syntactic dissociation between negation and quantification in the lexicon entry. Hence, there are distinct semantic denotations for each of the two syntactic features. The "some" operator simply denotes a generalized existential quantifier and is paired with corresponding meaning constructors, (14).

(14) **[some]** $\lambda P.\lambda Q.\exists x[P(x) \wedge Q(x)]$: $(\,(\mathrm{QUANT}\uparrow)_\sigma \multimap (\mathrm{QUANT}\uparrow)_\sigma) \multimap (\forall G\ (\,(\mathrm{QUANT}\uparrow)_\sigma \multimap G) \multimap G)$

The negation feature negates the proposition that is realized by its next higher f-structure node. Its semantic type is thus $< t, t >$. As explained above, selection of the correct type is ensured by a syntactic constraint forcing the availability of a subject in the f-structure containing the negatively quantified function. The meaning constructor of negation itself can therefore remain relatively simple. It merely asserts that negation is applied to the semantic contribution of its mother's f-structure and returns a semantic resource of the same kind, (15).

(15) **[neg]** $\lambda p.(\neg p)$: $(\forall H\ (\,(H\uparrow)_\sigma \multimap (H\uparrow)_\sigma)\,)$

The phrase structure rules in (11) and (12) are paired with an additional semantic premise, which I label **[xadj]**. This resource manages the integration of the initial XADJUNCT. Specifically, an entailment relation is assumed to hold between the meaning of an initial XADJUNCT and the meaning of its matrix clause. Hence, **[xadj]** asserts that if the initial constituent is true, the meaning of the whole clause is true. The denotation would have to be upgraded for more complicated examples (Dalrymple 2001: 352) but will suffice for the present purpose. The meaning constructors ensure the initial selection of the semantics of the clause-initial XADJUNCT as the antecedent argument, the subsequent application of the semantics of the matrix clause as the consequent argument, and the final derivation of a semantic resource that is also associated with the consequent, (16).

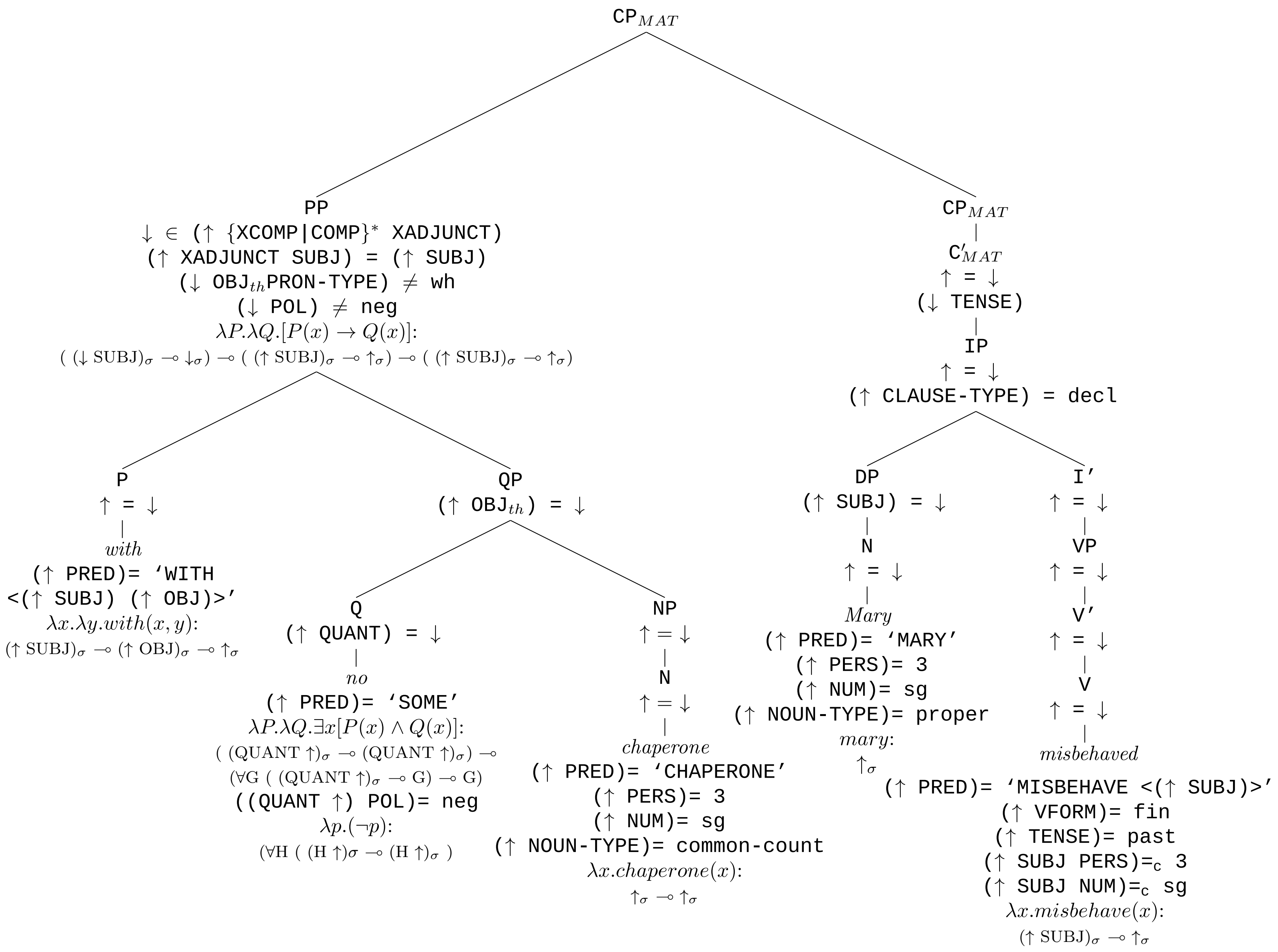
(16) **[xadj]** $\lambda P.\lambda Q.[P(x) \rightarrow Q(x)]$: $(\,(\downarrow \mathrm{SUBJ})_\sigma \multimap \downarrow_\sigma) \multimap (\,(\uparrow \mathrm{SUBJ})_\sigma \multimap \uparrow_\sigma) \multimap (\,(\uparrow \mathrm{SUBJ})_\sigma \multimap \uparrow_\sigma)$

The model can now provide syntactic parses and semantic compositions for a large number of sentences involving clause-initial PPs quantified by *no*. In the following paragraphs, I will illustrate the functioning of my model with cases of "optional inversion" after initial negative PPs. These structures have frequently been discussed in the literature (e.g., Lakoff 1991, Büring 2004, Haegeman 2012 to name but a few). The specific example I will use as an illustration is shown in (17) below.

(17) a. With no chaperone, Mary misbehaved.
b. With no chaperone did Mary misbehave.

An explicit analysis for the first sentence of this minimal pair is presented in (18). The analysis for the second sentence in the minimal pair follows immediately in (19).

(18) With no chaperone, Mary misbehaved.



$$f_1\left[\begin{array}{ll}
\text{PRED} & \text{‘MISBEHAVE}\langle(\text{SUBJ})\rangle\text{’}\\
\text{XADJUNCT} & \left\{ f_3\left[\begin{array}{ll}
\text{PRED} & \text{‘WITH}\langle(\text{SUBJ}),(\text{OBJ})\rangle\text{’}\\
\text{SUBJ} & [\quad]\\
\text{OBJ}_{th} & f_4\left[\begin{array}{ll}
\text{PRED} & \text{‘CHAPERONE’}\\
\text{QUANT} & f_5\left[\text{PRED}\ \ \text{‘SOME’}\right]\\
\text{PERS} & 3\\
\text{NUM} & \text{sg}\\
\text{NOUN-TYPE} & \text{common-count}\\
\text{POL} & \text{neg}
\end{array}\right]
\end{array}\right]\right\}\\
\text{SUBJ} & f_2\left[\begin{array}{ll}
\text{PRED} & \text{‘MARY’}\\
\text{PERS} & 3\\
\text{NUM} & \text{sg}\\
\text{NOUN-TYPE} & \text{proper}
\end{array}\right]\\
\text{TENSE} & \text{past}\\
\text{VFORM} & \text{fin}\\
\text{CLAUSE-TYPE} & \text{decl}
\end{array}\right]$$

**[with]** $\lambda x.\lambda y.with(x,y)$:

**[some]** $\lambda P.\lambda Q.\exists x[P(x) \wedge Q(x)]$:

**[chaperone]** $\lambda x.chaperone(x)$:

$$\dfrac{f_2 \multimap f_4 \multimap f_3 \qquad [s : f_2]^1}{\lambda y.with(s,y):} \multimap\varepsilon \qquad \dfrac{(f_4 \multimap f_4) \multimap (\forall G\ ((f_4 \multimap G) \multimap G) \qquad f_4 \multimap f_4}{\lambda Q.\exists x[chaperone(x) \wedge Q(x)]:} \multimap\varepsilon$$

$$\dfrac{f_4 \multimap f_3 \qquad (\forall G\ ((f_4 \multimap G) \multimap G)}{\exists x[chaperone(x) \wedge with(s,x)]:} \multimap\varepsilon$$

**[neg]** $\lambda p.(\neg p)$:

$$\dfrac{f_3 \qquad f_3 \multimap f_3}{\neg\exists x[chaperone(x) \wedge with(s,x)]:} \multimap\varepsilon$$

$$\dfrac{f_3}{\lambda y.\neg\exists x[chaperone(x) \wedge with(y,x)]:} \multimap\mathcal{I},1$$

**[xadj]** $\lambda P.\lambda Q.[P(x) \rightarrow Q(x)]$:

$$\dfrac{f_2 \multimap f_3 \qquad (f_2 \multimap f_3) \multimap (f_2 \multimap f_1) \multimap (f_2 \multimap f_1)}{\lambda Q.\lambda y.[\neg\exists x[chaperone(x) \wedge with(y,x)] \rightarrow Q(y)]:} \multimap\varepsilon$$

**[misbehave]** $\lambda x.misbehave(x)$:

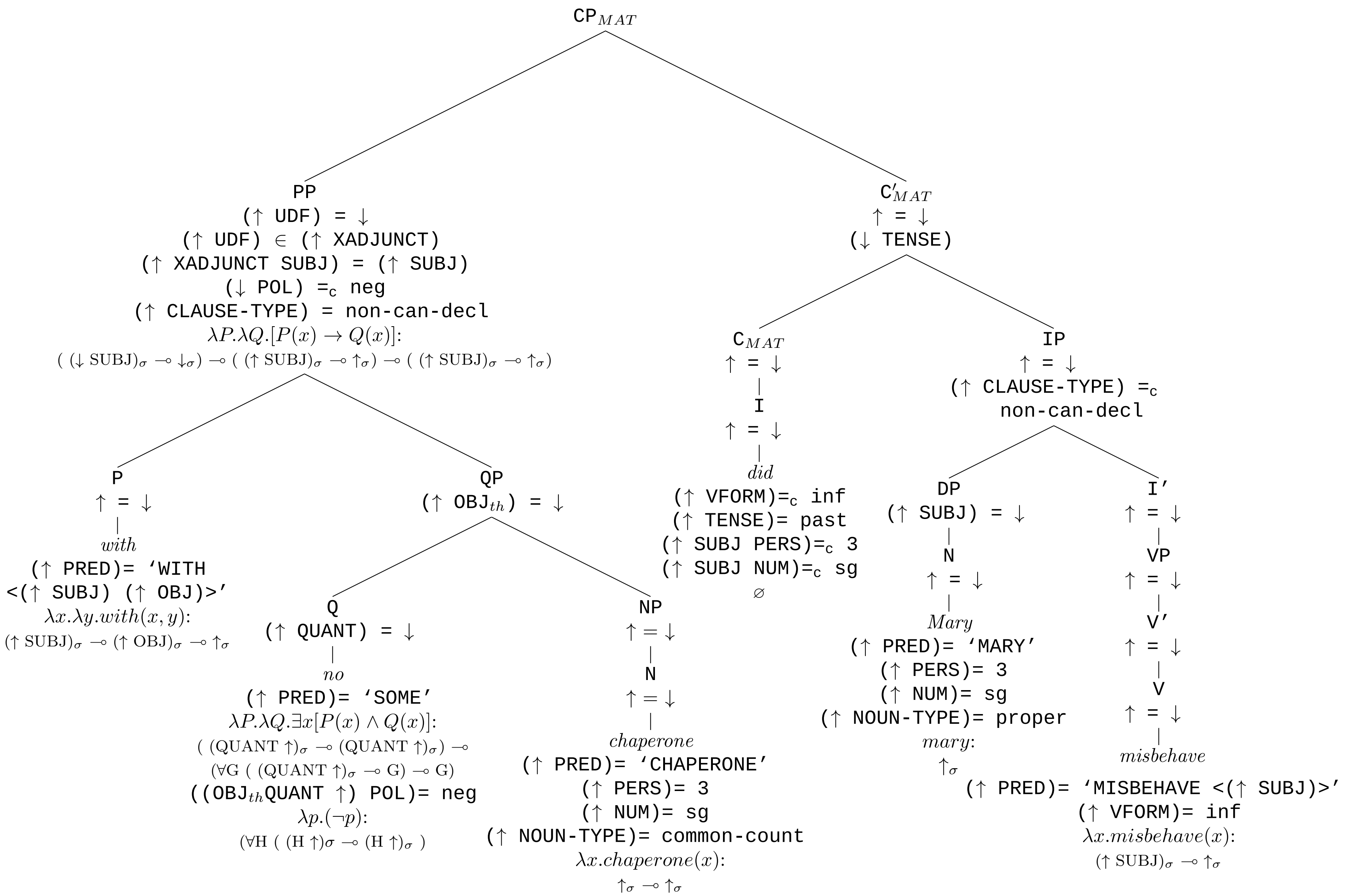
$$\dfrac{(f_2 \multimap f_1) \multimap (f_2 \multimap f_1) \qquad f_2 \multimap f_1}{\lambda y.[\neg\exists x[chaperone(x) \wedge with(y,x)] \rightarrow misbehave(y)]:} \multimap\varepsilon$$

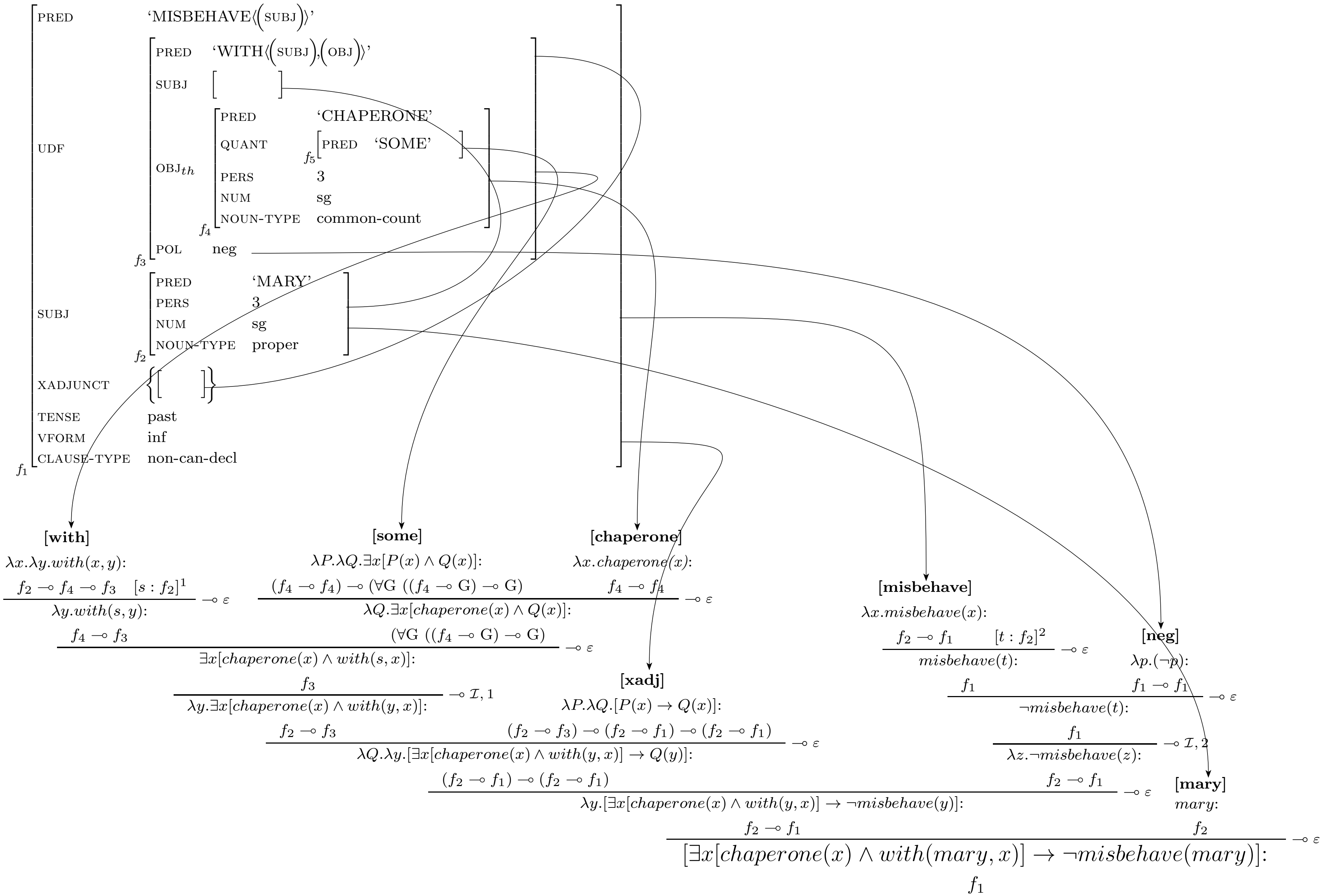
**[mary]** $mary$:

$$\dfrac{f_2 \multimap f_1 \qquad f_2}{[\neg\exists x[chaperone(x) \wedge with(mary,x)] \rightarrow misbehave(mary)]:} \multimap\varepsilon$$

$$f_1$$

(19) With no chaperone did Mary misbehave.




PRED ‘MISBEHAVE⟨(SUBJ)⟩’
UDF
PRED ‘WITH⟨(SUBJ),(OBJ)⟩’
SUBJ
OBJth
PRED ‘CHAPERONE’
QUANT
PRED ‘SOME’
PERS 3
NUM sg
NOUN-TYPE common-count
POL neg
SUBJ
PRED ‘MARY’
PERS 3
NUM sg
NOUN-TYPE proper
XADJUNCT
TENSE past
VFORM inf
CLAUSE-TYPE non-can-decl
f1 f2 f3 f4 f5
[with]
λx.λy.with(x, y):
f2 ⊸ f4 ⊸ f3
[s : f2]1
λy.with(s, y):
f4 ⊸ f3
[some]
λP.λQ.∃x[P(x) ∧ Q(x)]:
(f4 ⊸ f4) ⊸ (∀G ((f4 ⊸ G) ⊸ G)
[chaperone]
λx.chaperone(x):
f4 ⊸ f4
λQ.∃x[chaperone(x) ∧ Q(x)]:
(∀G ((f4 ⊸ G) ⊸ G)
∃x[chaperone(x) ∧ with(s, x)]:
f3
λy.∃x[chaperone(x) ∧ with(y, x)]:
f2 ⊸ f3
[xadj]
λP.λQ.[P(x) → Q(x)]:
(f2 ⊸ f3) ⊸ (f2 ⊸ f1) ⊸ (f2 ⊸ f1)
λQ.λy.[∃x[chaperone(x) ∧ with(y, x)] → Q(y)]:
(f2 ⊸ f1) ⊸ (f2 ⊸ f1)
[misbehave]
λx.misbehave(x):
f2 ⊸ f1
[t : f2]2
misbehave(t):
f1
[neg]
λp.(¬p):
f1 ⊸ f1
¬misbehave(t):
f1
λz.¬misbehave(z):
f2 ⊸ f1
λy.[∃x[chaperone(x) ∧ with(y, x)] → ¬misbehave(y)]:
f2 ⊸ f1
[mary]
mary:
f2
[∃x[chaperone(x) ∧ with(mary, x)] → ¬misbehave(mary)]:
f1
⊸ ε
⊸ I, 1
⊸ I, 2

The analysis of the first sentence works as follows. The clause-initial PP, *with no chaperone*, is parsed with the lexicon entry for *no* in (7). The first option for the placement of the negation feature is used, which puts negation in the f-structure corresponding to the range of the prepositional complement, *no chaperone*. The feature is licensed in this position because the XADJUNCT f-structure, in which the function it negates is embedded, contains a subject, *Mary*. As a consequence, the initial PP can only be parsed as a CP-adjunct by the rule in (11) - this rule explicitly bans a negation feature on the outer f-structure referring to the whole PP. Thus, a non-canonical declarative clause type feature is not introduced into the structure. Instead the clause-type attribute is valued as declarative in the IP-layer and the sentence thus shows canonical subject-verb word order. The model can generate exactly one solution for the syntactic parse.

The meaning of the sentence is computed as follows. First, the meaning constructors on the lexical items and phrase structure rules are instantiated relative to the f-structure constructed by the syntactic parse. The semantic composition then proceeds along the lines of the proof tree given. A proposition is composed for the phrase *with no chaperone*, "s was with some chaperone." The independent negation feature introduced by *no* must apply to this proposition based on its position in the f-structure. Hence, the polarity of the proposition is inverted, 'it is not the case that s was with a chaperone." The name "s" is later replaced by "Mary" via the formation of a variable by lambda abstraction over "s" and its subsequent saturation by functional application of "Mary." A second proposition is formed for the main clause *Mary misbehaved*, which is simply "Mary misbehaved." The two propositions are then put in a conditional relationship so that the final meaning of the whole sentence can be paraphrased as, "when Mary was not with a chaperone, she misbehaved." The logical formula representing this meaning is repeated in (20).

(20) $[\neg\exists x[chaperone(x) \wedge with(mary, x)] \rightarrow misbehave(mary)]$

This representation does in fact seem to be adequate for the sentence as it conforms to intuitions about its truth conditions (at least within the scope of first order logic, which abstracts away from issues such as time and events). To be precise, the formula assumes the sentence *With no chaperone, Mary misbehaved* to be false if Mary was not with a chaperone, but nevertheless behaved well, and to be true otherwise. The model provides one unique semantic composition for the sentence.

The second sentence is analyzed in the following way. As before, the initial PP, *with no chaperone*, is parsed with the lexicon entry for *no* in (7). However, the second option for the placement of the negation feature is now employed. As a consequence, negation is positioned within the f-structure that denotes the domain of the whole PP. The feature is licensed in this position because the matrix f-structure, which contains the negated function, also contains a subject, *Mary*. The initial PP can only be parsed as the specifier of CP with the phrase structure rule in (12), which expects the feature POL=neg on the PP-level, and not with the CP-adjunction rule in (11), which, conversely, prohibits the presence of such a feature in this place. Thus, a non-canonical declarative clause type feature is introduced into the structure. This clause type feature in turn triggers subject-auxiliary inversion by rule (61). The model produces one unique solution for the syntactic parse.

The semantic composition proceeds as shown in the proof tree. After instantiating the meaning constructors on the lexical items and phrase structure rules relative to the f-structure constructed by the syntax, the model constructs two propositions, which can be paraphrased as "s was with some chaperone" for the PP and "t misbehaved" for the main clause. The negation feature introduced by *no* negates the mother of the f-structure in which it is embedded. Hence, in the present case, it must scope over the main clause proposition, yielding "it was not the case that t misbehaved." New variables are created by lambda abstraction over "s" and "t." Those variables are bound under the same lambda operator as the two propositions are put into a conditional relationship and the second proposition undergoes functional application to the consequent argument. The variable is eventually saturated by functional application of "Mary." The final meaning of the sentence can be paraphrased as "when Mary was with a chaperone, then she did not misbehave." The logical formula representing this meaning is repeated in (21).

(21) $[\exists x[chaperone(x) \wedge with(mary, x)] \rightarrow \neg misbehave(mary)]$

There is, however, a second semantic formula that can be generated under the present assumptions. The model can form the proposition "when Mary was with a chaperone, she misbehaved." This proposition, too, represents the meaning assembled from the matrix f-structure. Hence, the meaning constructors of the negation feature allow for the application of this proposition to negation as well, yielding the meaning, "it was not the case that when Mary was with a chaperone, she misbehaved." In the proof tree above, the negation feature would be used not on the proposition "t misbehaved" but at the very last step in the derivation. The logical formula that would result from this proof is presented in (22).

(22) $\neg[\exists x[chaperone(x) \wedge with(mary, x)] \rightarrow misbehave(mary)]$

Interestingly, both formulas make the same truth-conditional claims for the scenario where it is in fact true that Mary was with a chaperone. In either case, the sentence is false if Mary was with a chaperone and misbehaved, and true if Mary was with a chaperone and did not misbehave. This is shown in the truth tables below, where p="Mary was with a chaperone" and q="Mary misbehaved."

| p | q | $p \rightarrow \neg q$ | $\neg (p \rightarrow q)$ |
|---|---|---|---|
| T | T | F | F |
| T | F | T | T |

Table A.1: Simplified truth tables for the formulas in (21) and (22)

The model thus generates two distinct semantic representations for the sentence, but both interpretations seem to be adequate as they correspond to intuitions about its truth conditions. It should be noted, however, that the two semantic formulas make different claims about the truth of the sentence in case Mary was not actually with a chaperone. The first formula would assert that the sentence is then false while the second would affirm its truth. I asked several native speakers of English if they thought the sentence *With no chaperone did Mary misbehave* was true or false if Mary was not actually with a chaperone. I did not get a coherent answer. This finding may show that the sentence presupposes that Mary was with a chaperone so that the antecedent proposition's falsity would lead to a presupposition failure. Interestingly, my model would then analyze presupposition failures as potential compositions of semantic representations that assert that the same sentence can be true or false at the same time. I shall not discuss in any more detail ramifications of this observation, but merely note that the fact that my model generates two distinct semantic formulas for sentences with subject-auxiliary inversion after negative propositional PPs may not be an unsurmountable challenge to the analysis proposed.

The model developed in the last sections can account for several additional facts about subject-auxiliary inversion after negative initial constituents. Firstly, it correctly requires obligatory inversion after certain negative phrases, such as the adverb *never*, (23), but also negative PPs such as *under no circumstances*, (24).

(23) a. Never does Mary eat apples.
  b. * Never, Mary eats apples.

(24) a. Under no circumstances did Mary want to be here tonight.
  b. * Under no circumstances, Mary wanted to be here tonight.

The mandatory nature of inversion in these examples follows directly from the fact that the phrases *never* and *under no circumstances* do not contain a subject function. Therefore, the negation feature cannot be placed within the range of the prepositional complement, but must be positioned in the f-structure referring to the whole initial phrase. Consequently, these constituents can only be parsed with the negation-sensitive phrase structure rules that place the initial phrases in the specifier of CP and introduce a non-canonical declarative clause type feature, making subject-auxiliary inversion obligatory (see rules (4) and (13)). My model thus incorporates the widely-held assumption that initial negative phrases that allow non-inversion with the subject must be "interpreted 'propositionally'" (Büring 2004: 6, and references cited therein).

My model also accounts correctly for possible and impossible extractions and resultant scope differences of negative initial constituents. The sentences in (25) illustrate the relevant observations for negative phrases without subject functions (that is, negative phrases that cannot be interpreted as propositions like those in the preceding two examples). Two positions of an underscore represent the negative phrase's interpretation site within the matrix and embedded clause respectively.

(25) a. *inversion and local extraction hence wide scope,* NEVER > SAY
  Never did Mary _ say that she had hit her husband.
  b. *inversion and long-distance extraction hence narrow scope,* SAY > NEVER
  *Never did Mary say that she had _ hit her husband.
  c. *no inversion and local extraction hence wide scope,* NEVER > SAY
  *Never, Mary _ said that she had hit her husband.
  d. *no inversion and long-distance extraction hence narrow scope,* SAY > NEVER
  *Never, Mary said that she had _ hit her husband.

The sentences in (25) all feature the clause-initial negative adverb *never* and must thus be parsed with the phrase structure rule in (4). Example (25a) shows subject-auxiliary inversion and the negative phrase is interpreted as an element of a local adjunct set. Since these are precisely the specifications required by (4), the sentence is correctly predicted to be grammatical. In contrast, example (25b), which also shows inversion, interprets the negative phrase as an element of the adjunct set of the embedded clause. The path required for such a long-distance dependency is not licensed by rule (4), and thus the sentence is correctly modeled as ungrammatical. Sentences (25c) and (25d) do not show inversion. The rule in (4), however, necessarily introduces a non-canonical declarative clause type feature, which triggers subject auxiliary inversion. Therefore my model cannot generate these sentences, as desired. Note that this means that negative phrases without a subject function simply cannot be extracted out of embedded clauses, irrespective of whether they occur with or without subject-auxiliary inversion (examples (25b) and (25d)). The thought itself would, of course, be perfectly fine, as shown by sentences such as *On Monday, Mary said that she had hit her husband*, which allows for the interpretation of the initial PP both in the matrix as well as in the embedded clause.

In contrast, my model works slightly differently for negative initial constituents with a subject function (that is, negative phrases that are interpreted as propositions unlike those in the preceding examples). Such phrases can in principle participate in a long-distance dependency as modeled by rule (11). Hence, the analogue of (25d) should be grammatical for initial phrases with subject functions under certain requirements. Specifically, the negative phrase must be adjoined to CP, subject-auxiliary inversion must not be triggered, and the negation feature must not be visible to the adjunction rule but must refer to a function more deeply embedded within the initial constituent, such as the complement of a preposition. An important side effect of the last condition is that the negation feature should not be able to scope over the proposition expressed by the subordinate clause. This aspect of my model does in fact seem to correspond to intuitions of the native speakers I consulted, as summarized by the example in (26).

(26) a. With no chaperone, Mary said that she misbehaved.

b. $say(mary, (\neg\exists x[chaperone(x) \wedge with(mary, x)] \rightarrow misbehave(mary)))$
Mary said, if she wasn't with a chaperone, she misbehaved.

c. *$say(mary, (\exists x[chaperone(x) \wedge with(mary, x)] \rightarrow \neg misbehave(mary)))$
*Mary said, no matter which chaperone she was with, she didn't misbehave.

The initial phrase *with no chaperone* in (26a) contains a negation feature introduced by the negative quantifier *no*. The negation feature must only refer to the prepositional complement and not to the whole PP because otherwise the constituent could neither be parsed as a CP-adjunct without subject-auxiliary inversion nor be interpreted as an element of the embedded clause. Consequently, the proposition expressed by the PP can apply to negation, "Mary wasn't with a chaperone," (26b) but the proposition expressed by the subordinate clause cannot, *"Mary didn't misbehave", (26c).

My model of Modern English clause structure can thus handle core aspects of the syntax and semantics of negative inversion structures.

# Appendix B

# A Brief Defense of Non-Structural Approaches to Verb Phrase Ellipsis

This appendix defends the claim that ellipsis sites do not contain unpronounced syntactic structures. Instead, the ellipsis resolution process described in the main part of the text can account for the interpretation of elided verb phrases. The appendix consists of two parts. First, I will outline how a non-structural approach can analyze sentences that are adduced as evidence of empty syntactic structure. Secondly, I will list arguments in favor of non-structural approaches.

**Countering Criticism of Non-Structural Approaches to Ellipsis**

There are certain kinds of VPE structures that are commonly believed to necessitate the presence of abstract syntactic structure in a presumed ellipsis site. Two of those construction types are the following.
(i) VPE with extraction. Elements can be extracted out of supposed ellipsis sites, including, among others, topicalized phrases or subjects of clausal complements of subject raising predicates. Example (1) involves the topicalized phrase *De Amicitia*, which is extracted across the board from the complement position of the preposition *at* in the source as well as from a putative ellipsis site in the target, as indicated with the arrows. Example (2) involves a source sentence with the verb *seem* and raising of a subject from its clausal complement to its local subject position, so that the following target VPE sentence is said to include an analogous kind of movement, again as shown by the arrows. Such examples are supposed to show that "VP-ellipsis sites allow for extraction via X$^0$-, A- and A'-, movement, thus corroborating the claim that they contain unpronounced syntactic structure" (Craenenbroeck and Merchant 2013: 707).

(1) But *De Amicitia* I could make a stab at _ , and could have Δ at any time in the last thirty-four years.

(Wallace Stegner (1987) *Crossing to Safety*, cited in: Merchant 2016: 7, ex. (19))

(2) John seems _ to be happy, but Mary doesn't Δ.

(adapted from: Craenenbroeck and Merchant 2013: 705, ex. (8a))

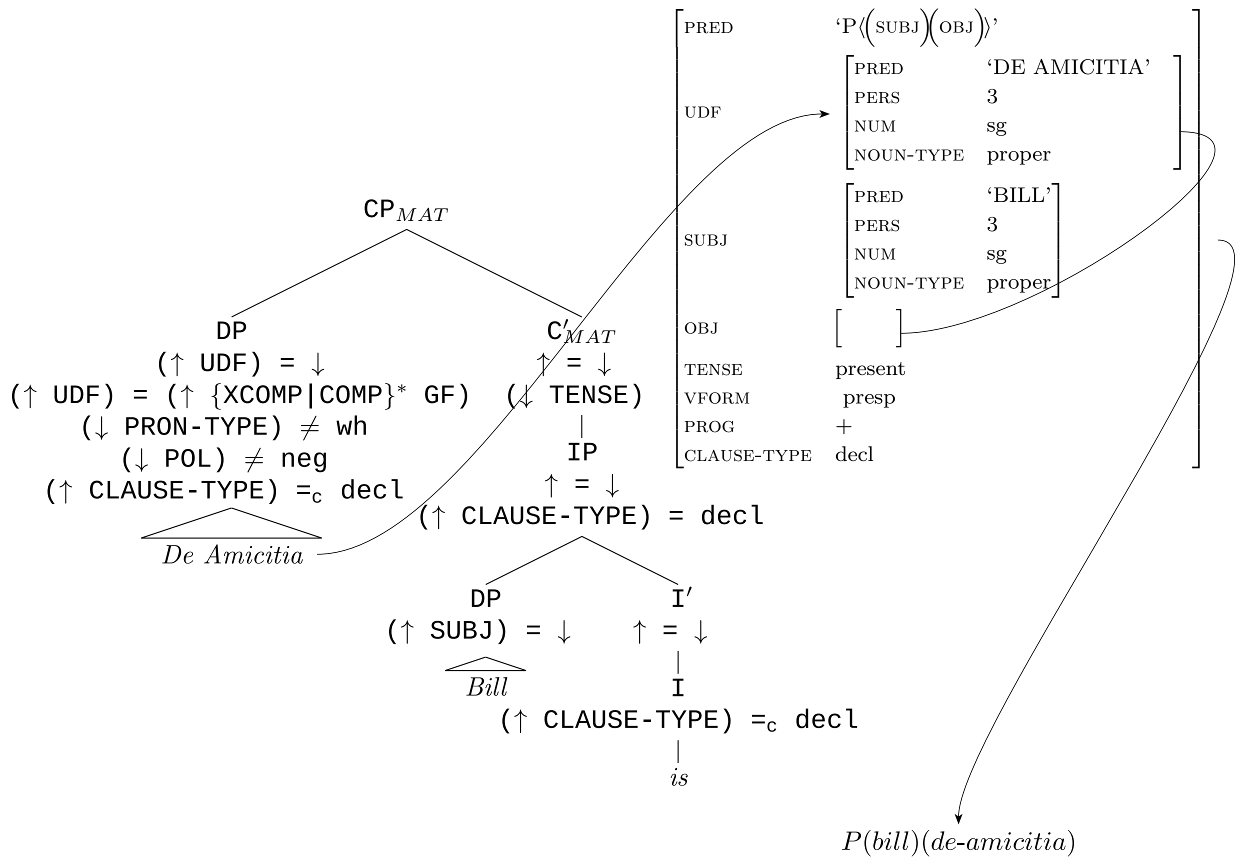
(ii) VPE in existential constructions. The existential construction, consisting of the expletive *there*, a form of the verb *to be* and an associate, shows agreement between the verb and the associate in number. This is true even if the associate is elided. In the example below, the source clause involves an existential construction with a plural associate, *many linguists*. The target sentence shows corresponding number agreement. This is taken to suggest that "the ellipsis site has to contain enough syntactic structure to host such an associate. Consequently, VPE [involves] [...] deletion of a fully-fledged verb phrase" (Aelbrecht 2009: 193).

(3) I didn't think there would be many linguists at the party.
But there { were / * was } Δ. (adapted from: Craenenbroeck and Merchant 2013: 704, ex. (3a))

If a non-structural, LFG model to verb phrase ellipsis is to be viable, it must be able to account for VPE structures with local extraction and associate number agreement ostensibly favoring structural approaches. I will show that the VPE mechanism explained in the main text of the thesis does in fact cope with these arguments quite effortlessly in conjunction with LFG's standard constraint-based means of handling issues of extraction and agreement. This demonstrates that the existence of such sentences does not constitute a knockout argument against non-structural approaches to VPE since non-structural approaches, such as the one proposed here, can account for them as well. Unfortunately, several linguistic phenomena that would be required for this purpose are not currently formally implemented in my model because they are not otherwise pertinent for the purposes of the chapter on possessive *have*. The following overview must thus necessarily be somewhat cursory and should merely be thought of as a proof of concept.

I will first consider the extraction argument. A straightforward example is provided by VPE structures with local topicalization. Such sentences can already be parsed by my model out of context without any further modifications. This is illustrated by the following example, *De Amicitia, Bill is* Δ.

(4) *De Amicitia*, Bill is Δ.




The parse functions largely as for ‘subject -verb’ VPE patterns, such as *Bill is* Δ. The only difference is that the present sentence also includes a clause-initial DP, *De Amicitia*, parsed with the rule in (7c), which places it in the specifier of CP and interprets its domain as an underspecified discourse function, UDF. The arrow leading from the clause-initial DP to its f-structure value illustrates this mapping. Moreover, the initial discourse function must be unified with another local or long-distance-dependent grammatical function (provided the initial phrase is not typed as a *wh*-element or as negative and that the clause type is declarative). In principle, *De Amicitia* could thus be unified with any grammatical function that yields a coherent parse. The current grammar fragment, however, makes only one subcategorization frame available that achieves this goal (though others are conceivable), namely an underspecified transitive predicate that requires a subject and an object function. Since the subject function is already present, the discourse function must be unified with the object. This unification is represented by a line ranging between the two respective feature structures. Again, the grammar produces one unique solution for this example. At this point, a relatively sophisticated interface between information structure and semantics would be required in order to produce a representation that adequately factors in the discursive role of argument fronting. In the absence of such an interface, I shall simply ignore the contribution from clause-initial discourse functions entirely and represent the meaning of such clauses as if they were ordinary transitives. The meaning of the VPE structure is thus given as some underspecified relation between *Bill* and *De Amicitia*.

The equational analysis will generate the relation that the resolution of the ellipsis requires. First, an appropriate source clause must be identified. Interestingly, a source clause for a target with topicalization is more natural if it itself includes a topicalized phrase. This is not further explicable without some implementation of the information structure-semantics interface, but underlines the fact that a certain degree of parallelism is required for ellipsis resolution. I will use the following toy discourse as an illustration.

(5) *(1) Identification of source and target*
*De Legibus*, Mary is reading. $\rightsquigarrow$ $read(mary, de\text{-}legibus)$
*De Amicitia*, Bill is $\Delta$. $\rightsquigarrow$ $P(bill)(de\text{-}amicitia)$

Secondly, one must establish parallel elements in the source and target clauses. Here, the subjects, *Mary* and *Bill*, and the topicalized objects, *De Legibus* and *De Amicitia*, constitute parallel elements. It is then possible to set up a parallelism equation, for which it is true that when the underspecified predicate, $P$, is applied to the interpretation of the parallel elements in the source, it will yield the interpretation of the source as a whole.

(6) *(2) Determination of the parallelism equation*
$P(mary, de\text{-}legibus) = read(mary, de\text{-}legibus)$

Using higher-order unification, one can now recover the relation $P$ from the parallelism equation.

(7) *(3) Solution of the parallelism equation for P*
$P(mary, de\text{-}legibus) = read(mary, de\text{-}legibus)$
$P(de\text{-}legibus) = \lambda x.read(x, de\text{-}legibus)$
$P = \lambda x.\lambda y.read(x, y)$

Finally, the substitution of this solution to the value of $P$ in the target sentence will generate the meaning of the ellipsis structure.

(8) *(4) Substitution of the solution of P for the underspecified predicate in the target*
$P = \lambda x.\lambda y.read(x, y)$
$P(bill)(de\text{-}amicitia) = \lambda x.\lambda y.read(x, y)(bill)(de\text{-}amicitia)$
$= read(bill, de\text{-}amicitia)$

The algorithm will work analogously for other cases of extraction in VPE structures, such as subject-to-subject raising constructions, etc. It follows that the existence of VPE structures with extraction out of a putative ellipsis site does not inescapably lead to the conclusion that they must contain unpronounced syntactic structure.

I will now turn to the argument from associate number agreement. It is substantially more difficult to explicitly describe the workings of the VPE resolution algorithm for existential constructions. For a start, a logical formula representing core aspects of the meaning of existential constructions is required for the application of the equational analysis to work. Unfortunately, the semantics of existential constructions are relatively ill-understood (for an overview, see McNally 2011). In particular, it is not entirely clear how to represent number information and if there should be semantic differences between sentences that contain existential expletives and those that show canonical subject-predicate structure. For simplicity's sake, I will make the assumptions that number information can be modeled as the cardinality of a set of individuals and that existential sentences can be represented as a set, or a conjunction of sets, of individuals bound under an existential operator. With this (inadequately simple) semantic formalization in place, one can thus represent the sentence *There were linguists at the party* with the formula in (9) below. It asserts that there is a set of entities that are both a linguist and at a party and that the cardinality of this set is larger than one.

(9) $\exists x[|linguist(x) \wedge at\text{-}party(x)| > 1]$

Furthermore, the syntactic structure of existential constructions must likewise be made explicit to allow for the generation of an f-structure and corresponding semantic formula for existential VPE sentences. Since the existential construction is in fact a very complex linguistic phenomenon - with unresolved issues ranging from co-occurrence restrictions of the expletive, associate and predicate over the identification of the predicate to the proper treatment of codas – I cannot provide a coherent syntactic analysis of them here without going beyond the scope of this chapter. A rough sketch of their syntax could read as follows: One could assume that a designated form of the verb *be* is used for existential constructions, which could be implemented, for instance, as a feature, `EXISTENTIAL=+`. A special phrase structure rule could map the associate onto the subject function in the f-structure while the expletive *there* in the canonical subject position could be represented as a subject feature, say `EXPL=+` (as in Sells 2005). With such formalizations in place, the set of subcategorization frames of an underspecified predicate, $P$, introduced by the VPE rules in (116) could be extended to cover VPE with existential constructions as well. Specifically, the rule would have to leave the subject `PRED` feature underspecified, introduce an appropriate number feature to agree with the verb, and it would also have to be sensitive to the presence of an `EXISTENTIAL` feature. With such a model, it would be possible to parse existential VPE structures, such as *There are* $\Delta$, to produce corresponding f-structures and semantic formulas along the following lines.

(10) There are $\Delta$.

$$\left[\begin{array}{ll}
\textsc{pred} & \text{'be}\big(\textsc{subj}\big)\text{'} \\
\textsc{subj} & \left[\begin{array}{ll}
\textsc{pred} & \text{'P'} \\
\textsc{pers} & 3 \\
\textsc{num} & \text{pl} \\
\textsc{noun-type} & \text{common-count} \\
\textsc{expl} & +
\end{array}\right] \\
\textsc{tense} & \text{present} \\
\textsc{vform} & \text{fin} \\
\textsc{clause-type} & \text{decl} \\
\textsc{existential} & +
\end{array}\right]
\qquad \exists x.[|P(x)| > 1]$$

Noteworthy in the f-structure above are the presence of the feature `EXISTENTIAL=+`, which regulates the applicability of the constraints associated specifically with existential constructions, the representation of the *there* expletive as `EXPL=+`, the underspecified predicate, $P$, heading the subject function, and, most importantly, the representation of the plural number information on the subject, `NUM=pl`, satisfying the agreement constraints on the finite verb.

The machinery thus sketched should allow for the resolution of existential VPE constructions under the equational analysis without any further modifications. To illustrate this, I will use the source and target sentences in the example discourse below.

(11) *(1) Identification of source and target*
Mary doesn't believe that there are linguists at the party. $\leadsto$
$\neg(believe(mary, (\exists x[|linguist(x) \wedge at\text{-}party(x)| > 1])))$
But there are $\Delta$. $\leadsto \exists x.[|P(x)| > 1]$

Next, one must identify parallel elements between source and target sentence. Here, the parallel elements are constituted only by the contribution of the structural expletive subject *there*, represented as the individual variables bound under an existential operator, and the subject's plural number feature, represented as the cardinality of a set of individuals. As before, the property that applies to the parallel elements in the source can be equated with the underspecified predicate, $P$, in the parallelism equation.

(12) *(2) Determination of the parallelism equation*
$\neg(believe(mary, (\exists x[|P(x)| > 1]))) = \neg(believe(mary, (\exists x[|linguist(x) \wedge at\text{-}party(x)| > 1])))$

This equation can be solved for $P$.

(13) *(3) Solution of the parallelism equation for P*
$\neg(believe(mary, (\exists x[|P(x)| > 1]))) = \neg(believe(mary, (\exists x[|linguist(x) \wedge at\text{-}party(x)| > 1])))$
$\exists x[|P(x)| > 1] = \exists x[|linguist(x) \wedge at\text{-}party(x)| > 1]$
$|P(x)| > 1 = |linguist(x) \wedge at\text{-}party(x)| > 1$
$P(x) = linguist(x) \wedge at\text{-}party(x)$
$P = \lambda x.linguist(x) \wedge at\text{-}party(x)$

The solution for $P$ can then be substituted for the underspecified predicate in the VPE sentence. In this example, the presence of the existential operator in the denotation of the VPE structure leads to existential closure over the individual variables. This yields a semantic representation for the ellipsis structure that correctly corresponds to intuitions about its meaning in the above context.

(14) *(4) Substitution of the solution of P for the underspecified predicate in the target*
$P = \lambda x.linguist(x) \wedge at\text{-}party(x)$
$\exists x.[|P(x)| > 1] = \exists x.[|linguist(x) \wedge at\text{-}party(x)| > 1]$

I will now go through the same example but include a number mismatch between the source and target sentences, as in the following discourse.

(15) Mary doesn't believe that there is a linguist at the party.
?? But there are Δ.

First of all, several of my native speaker informants marginally accept the above discourse contrary to the expectation that its number mismatch should rule out the VPE structure entirely. I therefore mark the VPE structure with two question marks rather than as ungrammatical. However, I assume that the marginal acceptability of this discourse is enforced by pragmatic identifications of parallel elements and that the interpretation of the VPE sentence should be impossible if parallel elements were established on syntactic grounds alone, as is the case under my simplistic assumptions. Now, my algorithm would determine that the number information is not parallel between the source and target sentences in this example, and it should consequently not be included in the term including the underspecified predicate, $P$, on the left hand side of the parallelism equation, (16a). As a consequence, the solution of $P$ will contain the number information of the source clause, (16b). The substitution of that solution of $P$ for the underspecified property in the target sentence will then result in a semantic representation that inconsistently asserts that the number of elements in the set of linguists at a party is both equal to and greater than one, (16c). The version of the equational analysis used here will therefore rule out such discourses, as desired.

(16) a. $\neg(believe(mary, (\exists x[P(x)]))) = \neg(believe(mary, (\exists x[|linguist(x) \wedge at\text{-}party(x)| = 1])))$
b. $P = \lambda x.|linguist(x) \wedge at\text{-}party(x)| = 1$
c. $* \exists x.[||linguist(x) \wedge at\text{-}party(x)| = 1| > 1]$

The fact that the equational analysis can successfully compute the meaning of existential VPE sentences and consequently rule out mismatches in associate-verb number agreement between source and target clauses demonstrates that these constructions cannot serve as an unequivocal piece of evidence for the justification of abstract syntactic structure in putative ellipsis sites.

**Positive Support for Non-Structural Approaches to Ellipsis**

So far I have defended the non-structural approach to VPE used here against scholars who insist that such a conceptualization be unviable. I would like to conclude this section by listing four positive arguments that suggest that a non-structural approach to VPE is in fact preferable to a structural one.

Firstly, structural approaches struggle to explicitly rule out cases where the unpronounced syntactic material in the ellipsis bears no resemblance to the overt material of the antecedent clause. Such a case is illustrated in the example below, which shows unpronounced syntactic structure in angled brackets, and which cannot be interpreted with this material.

(17) * Potatoes, I like. But tomatoes$_i$, I don't <buy $t_i$ at our local store>.

Structural approaches commonly posit conditions on identity between the antecedent and the clause containing the ellipsis site. For example, it has been suggested that a mutual entailment relationship must hold between antecedent and elided verb phrase (e.g., the notion of e-GIVEN in Merchant 2001). Alternatively (or additionally), new words or lexical morphemes that are not found in the antecedent are banned from entering the derivation of the ellipsis (e.g., recoverability of sluicing in Chung 2006). However, such identity conditions are usually formulated as prerequisites for the calculation of the meaning of missing material during parsing, not as constraints on the introduction of syntactic material during production. Hence, it remains unclear by which specific mechanism the unpronounced syntactic material is made sensitive to these conditions. Or put differently, I do not see how a condition on identity actually leads to the presence of the right kind of syntactic structure. Is the syntactic structure somehow copied from the antecedent into the ellipsis? If so, the identity conditions seem to be relatively redundant. Does the syntax generate a large (or possibly infinite) number of verb phrases that can potentially be elided and that will then be filtered out according to the identity condition? In this case, the computational burden imposed by the syntax of VPE would appear to be implausibly heavy. The analysis adopted here does not posit empty syntactic structure and therefore does not need to explain the establishment of a syntactic relationship between antecedent and ellipsis either.

Secondly, there are several kinds of form mismatches between the syntax of the source clause and the syntax that could be reconstructed from the VPE target clause. These include inconsistencies in verbal inflections as in (18) (e.g., Quirk et al. 1972, Warner 1985, Potsdam 1997), voice as in (19) (e.g., Dalrymple et al. 1991, Fiengo and May 1994) and negative polarity items as in (20) (e.g., Sag 1976).

(18) *inflectional morphology mismatch: source - past participle, target - infinitive*
   a. Mary has eaten a sandwich, and Bill will $\Delta$, too.
   b. * Mary has eaten a sandwich, and Bill will eaten a sandwich, too.

(19) *voice mismatch: source - passive, target - active*
   a. A lot of this material can be presented in a fairly informal and accessible fashion, and often I do $\Delta$. (Noam Chomsky (1982) *Noam Chomsky on the Generative Enterprise*, cited in: Dalrymple et al. 1991: 440, (ex. 59a))
   b. * A lot of this material can be presented in a fairly informal and accessible fashion, and often I do be presented (this material) in a fairly informal and accessible fashion.

(20) *NPI mismatch: source - NPI, target - no NPI*
   a. John doesn't see anyone, but Bill does $\Delta$.
   (adapted from: Sag 1976: 157, (ex. 2.3.39))
   b. * John doesn't see anyone, but Bill does see anyone.

Structural approaches are challenged by such mismatches because the empty syntactic material in the supposed ellipsis site is assumed to be largely identical to the corresponding overt material in the antecedent clause (by means of some identity condition, as just explained). Structural approaches thus require ad-hoc repairs: Certain morphemes or features, say verbal inflections, can be ignored for the satisfaction of identity conditions; voice mismatches can be accounted for by postulating a voice projection that is outside of the elided verb phrase (e.g., Merchant 2008*a*) etc. Admittedly, my account's specific predictions regarding form mismatches will greatly depend on (i) the nature of the semantic representations used for the equational analysis and (ii) how exactly parallel elements are defined. But other than that, it does not require any further adjustments to deal with form mismatches of the above types or other kinds.

Thirdly, there are several cases of ambiguous interpretations for the meaning of the VPE structure. These include associations with different antecedent clauses as in (21), examples of strict and sloppy identity with varying degrees of complexity as in (22), or Hirschbühler's Canadian Flag sentences as in (23).

(21) Mary will sell a handbag. If George comes to watch, Sarah will $\Delta$, too .
   a. $\Delta$=sell a handbag, b. $\Delta$=come to watch

(22) Bill believed that he loved his wife, and Harry did $\Delta$, too.
   a. $\Delta$=believe that Bill loved Bill's wife.
   b. $\Delta$=believe that Harry loved Harry's wife.
   c. $\Delta$=believe that Harry loved Bill's wife.
   d. *$\Delta$=believe that Bill loved Harry's wife.
   (adapted from: Dalrymple et al. 1991: 447, who cite Dahl 1974 as a source)

(23) A Canadian flag is in front of every building, and an American flag is $\Delta$, too.
   a. $\exists$ Canadian flag $>$ $\forall$ buildings; $\exists$ American flag $>$ $\forall$ buildings
   b. *$\exists$ Canadian flag $>$ $\forall$ buildings; $\forall$ buildings $>$ $\exists$ American flag
   c. $\forall$ buildings $>$ $\exists$ Canadian flag; $\forall$ buildings $>$ $\exists$ American flag (implausible, but possible)
   d. *$\forall$ buildings $>$ $\exists$ Canadian flag; $\exists$ American flag $>$ $\forall$ buildings

Depending on what structural theory one adopts exactly (for a relatively successful approach, see Fox 2000), the analysis of these interpretative ambiguities may involve a large number of stipulations (e.g., economy conditions), may refer to the fact of ellipsis itself as a contextual condition or, if they are too naive, may lead to plainly wrong empirical predictions. In contrast, non-structural approaches similar in spirit to what has been proposed here can correctly predict these and various other ambiguities with great elegance and with reference to virtually nothing but general principles of semantic composition (for strict and sloppy identity, see Shieber et al. 1996; for Canadian flag sentences, see Asudeh and Crouch 2002, and references therein). I take it that this speaks in favor of such analyses.

Fourthly, most structural approaches should predict long-distance extraction out of putative ellipsis sites to be grammatical, contrary to fact, as shown in (24).

(24) a. *local dependency - VPE possible*
The island of Taiwan, typhoons have [frequently destroyed].
The island of Haiti, hurricanes have $\Delta$.

b. *long-distance dependency - VPE impossible*
The island of Taiwan, Bill has [said that thunderstorms have frequently destroyed]
* The island of Haiti, John has $\Delta$.

Proponents of structural approaches to VPE often respond to this observation in one of three ways. (i) They may acknowledge the impossibility of long-distance extraction only with respect to syntactic islands. An example is given below.

(25) *impossible VPE long-distance extraction out of an adjunct island*

a. * Ben will be mad if Abby talks to Mr Ryberg, and guess who CHUCK will.

b. * Ben will be mad if Abby talks to Mr Ryberg, and guess who CHUCK will be mad if Abby talks to.
(adapted from: Craenenbroeck and Merchant 2013: 707, ex. (12b), (13b))

It is then possible to rule out a subset of the offending structures based on independent syntactic principles in line with the structural paradigm. The obvious problem here consists in the unexplained residue of data (e.g., (24b)). (ii) They may posit a condition that generalizes a ban on eliding less of a structure than possible (e.g., the principle of MaxElide in Merchant 2008*b*: 141). This solution seems questionable because it remains unclear how exactly such a condition would come about. (iii) They may argue that long-distance dependencies of the relevant kind are not actually ungrammatical and are open to amelioration and repair (e.g., Fox and Lasnik 2003). The examples below present supposedly grammatical examples of long-distance extractions out of ellipsis sites. The grammaticality judgments are shown as in the original publications.

(26) a. $\text{ABBEY}_1$ said $\text{she}_1$ took GREEK, but I don't remember what language $\text{BETH}_2$ did <say $\text{she}_2$ took>.
(adapted from: Merchant 2008*b*: 140, ex. (13f))

b. ?? I know which book John said that Mary read, but YOU don't know which one he did.
(adapted from: Fox and Lasnik 2003: 151, ex. (38b))

However, my native speaker informants agree that those sentences should be regarded as ungrammatical. Hence, even if certain factors can ameliorate a VPE sentence to some degree in comparison to its non-elided counterpart, it should still not be part of the language defined by a grammar model of Modern English. Under the non-structural approach proposed above, a ban on long-distance extraction in VPE structures follows directly. The reason is that the determination of the parallelism equation with two parallel elements (say a topicalized object and a subject) requires a term on the left hand side that expresses a relation between those two elements. The two elements must therefore be in a local relationship (or involve structure sharing with a more deeply embedded element).