Fall 12-14-2009

Antisymmetry and the Conservation of C-Command: Scrambling and Phrase Structure in Synchronic and Diachronic Perspective

Joel C. Wallenberg
University of Pennsylvania, joel.wallenberg@gmail.com

Antisymmetry and the Conservation of C-Command: Scrambling and Phrase Structure in Synchronic and Diachronic Perspective

Abstract
Holmberg's Generalization (Holmberg 1986) was originally stated to describe the "object shift" phenomena found in the modern Scandinavian languages. This dissertation argues that object shift is merely a subcase of scrambling, a type of adjunction, and that Holmberg's Generalization is a subcase of a universal constraint, the "Generalized Holmberg Constraint" (GHC), which prohibits leftward scrambling across c-commanding functional heads. The existence of such a constraint turns out to have ramifications far beyond the analysis of scrambling itself, and the predictions it makes ultimately form an extended argument in favor of a universal antisymmetric approach to phrase structure (Kayne 1994).

The most important evidence for the GHC comes from diachronic data. The study presents quantitative data from the history of Yiddish and English to show that, in cases where a language undergoes major changes in its clause structure, the GHC remains an active and stable constraint in the language, indicating its status as a universal. Once a phrase structure change begins, the resulting variation within a single speech community, and even within individuals, immediately shows the effect of the GHC on scrambling.

The latter portion of the study argues that the GHC is not merely a constraint on scrambling, but rather a much more general constraint on the way syntactic computations progress, the "Conservation of C-Command." The Conservation of C-Command finds a natural cross-linguistic formulation only if we adopt an antisymmetric approach to languages with head-final phrase structures. This approach turns out to have consequences for a variety of other problems of syntactic analysis, including the West Germanic Verb (Projection) Raising construction and Heavy NP Shift.

This dissertation accounts for the typology of scrambling found in the world's languages and during periods of language change, and shows that the way in which scrambling is constrained provides insight into basic properties of phrase structure. In addition, it constitutes an extended argument for the autonomy of syntax: while prosodic and pragmatic considerations favor leftward scrambling in a number of contexts, a language's inventory of functional heads puts a strict upper bound on whether scrambling can respond to these considerations.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Linguistics

First Advisor
Anthony S. Kroch

Keywords
syntax, scrambling, language change, historical, Germanic

Subject Categories
Comparative and Historical Linguistics | Syntax

ANTISYMMETRY AND THE CONSERVATION OF C-COMMAND: SCRAMBLING AND PHRASE STRUCTURE IN SYNCHRONIC AND DIACHRONIC PERSPECTIVE

Joel C. Wallenberg

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2009

Supervisor of Dissertation

_____

Anthony S. Kroch, Professor of Linguistics

Graduate Group Chairperson

_____

Eugene Buckley, Associate Professor of Linguistics

Dissertation Committee

Anthony S. Kroch, Professor of Linguistics
Donald Ringe, Professor of Linguistics, Department Chair
Charles Yang, Associate Professor of Linguistics and Computer Science
Höskuldur Thráinsson, Professor of Icelandic, University of Iceland (Háskóli Íslands)

Antisymmetry and the Conservation of C-Command: scrambling and phrase structure in synchronic and diachronic perspective

Acknowledgements

I would first and foremost like to thank my wife, Caitlin Light, whose intellectual and emotional support made this possible in any number of ways.

Anthony Kroch has been an outstanding advisor and mentor to me, both in terms of my development as a linguist (and the development of many specific analyses that appear here and elsewhere), and in terms of my general thinking about many aspects of life. As Immanuel Kant once wrote of David Hume, Tony "woke me from my dogmatic slumber".

A great deal of credit for my intellectual development as a linguist goes to the faculty and students of the University of Pennsylvania. But before I give in to the urge to list them all (which I'm fighting), I would like to thank my dissertation committee, most of whom have been with me on this journey since the proposal stage: Don Ringe, Höskuldur Thráinsson, and Charles Yang. I would like to single out Don Ringe as a second mentor to me, and to thank him for convincing me that I'm as much an historical linguist at heart as I am a syntactician; I hope I can live up to that! He always shows up with clear and exceptionally well-informed thinking on every subject he addresses. Similarly, thank you to Bill Labov and Gillian Sankoff for convincing me that variation and change always go hand in hand, and I promise to keep variation and language contact in mind when addressing any linguistic problem. Thanks also to Julie Legate for a number of comments on various stages of this project.

Many thanks go to Beatrice Santorini, for a great deal of insightful discussion, especially concerning Early and Modern Yiddish data, as well as much help in working with parsed corpora in general. I also want to thank her for being an excellent friend and helping me address a number of difficult issues late in the process of writing this dissertation.

The students (many now faculty) at Penn who have influenced me are far too many to name individually, but here are a small few of those to whom I owe gratitude. Thank you Marjorie Pak for many interesting discussions, but especially for suggesting that I should look at Kayne's Antisymmetry Hypothesis! Five minutes of discussion of the issue with you was enough to change the course of my research for years. Thank you to my cohort at Penn – congratulations guys, we all finished! – and especially to Tatjana Scheffler and Maya Ravindranath for being close friends and colleagues from the very beginning of this process. I would also like to thank all of the various members of the Dr. Cardona Happy Hour group: Augustin Speyer, Josef Fruewald, Aviad Eilam, Jonathan Gress-Wright, Ron Kim from time to time, whoever else I've forgotten, and especially Dr. Jean-Francois Mondon. Thank you to Tom McFadden for being my "buddy" upon arrival, and to Sandhya Sundaresan for being another great friend over the years. Also thanks goes to Mark Dredze for, among other things, helping me get a foot in the door of machine learning and computational linguistics more generally.

I'd like to thank Theresa Biberauer and Ian Robert for a number of helpful conversations regarding the theoretical bent of this dissertation, and for much needed moral support at key moments.

Thank you to Arnold Zwicky and Paul Kiparsky, who have both been major influences on my linguistic thinking and good friends since my undergraduate days; long may it continue!

I would like to thank my parents and the rest of my family for putting up with many emotional telephone calls during this process, and for general, unfailing support. I would also like to thank Lauren Hall-Lew and other members of the Dissertation Cops group for their extremely important emotional support during the period in which I was actually writing this thesis. Similarly, Central Philadelphia Monthly Meeting has been an important support to me during the last two years of writing and various life-changes. Also, William Shakespeare for stress relief, and Sherlock Holmes, without whom I wouldn't have started thinking about object shift in English.

ABSTRACT

ANTISYMMETRY AND THE CONSERVATION OF C-COMMAND: SCRAMBLING AND PHRASE STRUCTURE IN SYNCHRONIC AND DIACHRONIC PERSPECTIVE

Joel C. Wallenberg

Supervisor: Anthony S. Kroch

Holmberg's Generalization (Holmberg 1986) was originally stated to describe the "object shift" phenomena found in the modern Scandinavian languages. This dissertation argues that object shift is merely a subcase of scrambling, a type of adjunction, and that Holmberg's Generalization is a subcase of a universal constraint, the "Generalized Holmberg Constraint" (GHC), which prohibits leftward scrambling across c-commanding functional heads. The existence of such a constraint turns out to have ramifications far beyond the analysis of scrambling itself, and the predictions it makes ultimately form an extended argument in favor of a universal antisymmetric approach to phrase structure (Kayne 1994).

The most important evidence for the GHC comes from diachronic data. The study presents quantitative data from the history of Yiddish and English to show that, in cases where a language undergoes major changes in its clause structure, the GHC remains an

active and stable constraint in the language, indicating its status as a universal. Once a phrase structure change begins, the resulting variation within a single speech community, and even within individuals, immediately shows the effect of the GHC on scrambling.

The latter portion of the study argues that the GHC is not merely a constraint on scrambling, but rather a much more general constraint on the way syntactic computations progress, the "Conservation of C-Command." The Conservation of C-Command finds a natural cross-linguistic formulation only if we adopt an antisymmetric approach to languages with head-final phrase structures. This approach turns out to have consequences for a variety of other problems of syntactic analysis, including the West Germanic Verb (Projection) Raising construction and Heavy NP Shift.

This dissertation accounts for the typology of scrambling found in the world's languages and during periods of language change, and shows that the way in which scrambling is constrained provides insight into basic properties of phrase structure. In addition, it constitutes an extended argument for the autonomy of syntax: while prosodic and pragmatic considerations favor leftward scrambling in a number of contexts, a language's inventory of functional heads puts a strict upper bound on whether scrambling can respond to these considerations.

Contents

3 Scrambling and Phrase Structure Change in Yiddish

4 Early Modern English Scrambling/Object Shift and the V-to-T Movement Parameter

5 Scrambling and the Antisymmetry Hypothesis

Chapter 1

Introduction

1.1 A Unified Analysis of Object Shift and Scrambling Cross-linguistically

Holmberg's Generalization (first stated in Holmberg 1986) was originally stated to describe the "object shift" phenomena found in the modern Scandinavian (North Germanic) languages. Object shift is a type of leftward DP scrambling which moves objects leftward across various clausal constituents, but crucially does not move objects across the main verb of a clause. For this reason, objects can only be moved over most clausal consituents (e.g., negation, or other left-vP-adjoined adjuncts) if the lexical verb also moves leftward. In most of the mainland Scandinavian languages/dialects, which are VO and V2 but do not have general V-to-T movement, this means that object shift over elements like negation can only take place in main clauses without auxiliaries, in which the finite lexical verb moves to C. In Icelandic and Faroese, which have general V-to-T movement, objects can shift over constituents to the left of vP in both matrix and subordinate clauses, by following the verb in its movement to Tense. In none of these languages can objects scramble leftward across a nonfinite verb. The basic pattern is shown in the examples below for Swedish and Icelandic (cf. Holmberg 1986, Holmberg

1999, Hellan & Platzack 1995, Thráinsson 2001, and references therein, for further discussion of the basic facts in Scandinavian).

(1) a. Varför läste$_i$ studenterna **den$_j$** inte t_i t_j?
why read students-the it not
"Why didn't the students read it?"

b. * Varför har studenterna **den$_i$** inte läst t_i?
why have students-the it not read

c. * Hon frågade varför studenterna **den$_i$** inte läste?
She asked why students-the it not read
"She asked why the students didn't read the books."

(2) a. Af hverju lásu$_i$ nemendurnir **bækurnar$_j$** ekki t_i t_j
for what read students-the books-the not
"Why didn't the students read the books?"

b. * Af hverju hafa nemendurnir **bækurnar$_i$** ekki lesið t_i?
for what have students-the books-the not read

c. Hún spurði af hverju stúdentarnir læsu$_i$ **bækurnar$_j$** ekki t_i t_j
She asked for what students-the read books-the not
"She asked why the students didn't read the books."
(Thráinsson 2001: 152)

Speakers of all of the Scandinavian languages allow (and frequently require) their unstressed, weak object pronouns to undergo overt object shift when their governing lexical verb moves left, but the languages differ as to whether full DP objects can scramble as well. I take the ability to scramble DPs to be an independent parameter on which languages obviously can differ. However, even in Icelandic, which optionally scrambles definite DPs leftward, an object cannot be scrambled across a verb. The observation that object shift is parasitic on verb-movement in Scandinavian is due to Holmberg (1986: 165), and is known as "Holmberg's Generalization".

Holmberg's Generalization, as he put it in Holmberg (1997: 208) stated, that "Object Shift is blocked by any phonologically visible category preceding/c-commanding the object position within VP". This statement of Holmberg's Generalization has the following drawbacks:

1. Objects cannot scramble past Tense/Infl or Comp in Scandinavian either, and these elements are outside of the VP.
2. The statement refers both to phonology and syntax; can this be simplified?
3. The statement is not clear about the relationship between linear precedence and c-command; does object shift tell us anything about this relationship?
4. The statement only applies to Scandinavian "Object Shift". Is there a wider phenomenon to be investigated here?

This dissertation will address all of these points, as well as a number of others which arise along the way. And in particular, with regard to the question in the last point, this dissertation gives a resounding "yes". There is, as we will see, a more general statement of the constraint at work in Holmberg's Generalization which can extend to other languages showing scrambling phenomena, e.g. German scrambling. However, Holmberg's Generalization, as stated above, cannot extend to cover the case of German scrambling, for example, since German is an OV language unlike the Scandinavian languages: for Scandinavian, Holmberg is able to say "preceding/c-commanding", but for under standard analyses of OV Germanic, the verb follows the VP but still c-commands it, so precedence and c-command must be stated separately. Furthermore, as I show

below, scrambling is blocked by elements outside of the VP/*v*P as well, so something must be added to Holmberg's Generalization.

The goal of this dissertation is to show that Holmberg's Generalization is not merely a fact about Scandinavian syntax, but rather a cross-linguistic constraint on leftward scrambling processes, and plausibly a language universal. The constraint can be stated in the following, more general way:

1st Version:

The Generalized Holmberg Constraint (GHC):

Scrambling moves phrases to left-phrasally-adjoined positions,[1] and may not cross a c-commanding head on the left in which a morpheme has been merged (i.e. internally, by head-movement, or externally, by substitution).

Furthermore, the constraint is diachronically stable, even as other aspects of a language's phrase structure are changing. Changes in the headedness of functional projections, for instance, interact with the Generalized Holmberg Constraint to produce different surface patterns, but the constraint remains constant. In this way, the diachronic aspect of this dissertation is perhaps less about language change, as it is about what doesn't change.

The first part of the dissertation, from here until Chapter 5, will make a case for the existence of the GHC, and will do so under a classical phrase structure. By "classical", I mean the approach to phrase structure that was generally assumed prior to

[1] Actually, the first clause of this constraint is derivable from Baltin's (1982) "Like Attracts Like Constraint", but I have repeated it here for clarity.

Kayne (1994) and Chomsky (1995), in which the headedness of phrases could vary independently of the hierarchical structure of phrases (the Head Parameter governs this variation). Right-headed structures are the mirror image of left-headed structures under this view; same hierarchy, different precedence relations between head and complement. Under this view, there can also be rightward or leftward movement operations, though scrambling is always assumed to be leftward.

This approach to headedness is assumed until Chapter 5 for ease of exposition only, and will ultimately be rejected in that chapter. At that point, I will develop an antisymmetric approach to head-final languages, borrowing heavily from the ideas of Kayne (1994), Chomsky (1995), and Biberauer (2003a). The GHC will be restated at that point in its final version, as shown below:

2nd (and Final) Version:

Conservation of C-Command:

Adjunction cannot subtract a c-command relation holding between a head and a non-head.

1.2 Organization of this Thesis

This dissertation argues that object shift is merely a subcase of scrambling, a type of adjunction (Saito 1985, Webelhuth 1989, Lee 1993, *inter alia*), and that Holmberg's Generalization is a subcase of a universal constraint, the "Generalized Holmberg Constraint" (GHC) or later, the "Conservation of C-Command", which prohibits leftward scrambling across c-commanding functional heads. The first part of the dissertation focuses on the evidence for the existence of the GHC as a generalization about scrambling. The latter portion of the study argues that the GHC is not merely a constraint on a single construction (scrambling), but rather that it is a much more general constraint on the way syntactic computations progress, the "Conservation of C-Command". In fact, while the dissertation initially describes the GHC in terms of classical X-bar theory, the GHC finds a natural cross-linguistic formulation in an antisymmetric approach to the traditional "headedness parameter." The empirical fact of the GHC's existence helps us decide between alternative theories of basic phrase structure.

In Chapter 2 I discuss the typology of scrambling phenomena inside and outside of Germanic, including OV and VO languages, Tense-final and Tense-medial languages, and Comp-final and Comp-medial languages. The chapter shows that the upper bound on the number of landing sites a language can make available to scrambled constituents is determined by the set of functional heads to the left of the constituent's base position. Particular detail is given to scrambling in German and Yiddish, and the lack of leftward scrambling across functional heads also provides a strong argument that Yiddish is underlyingly an OV language, *contra* Diesing (1997).

Perhaps the most important evidence for the GHC comes from diachronic data. This study presents quantitative data from a number of corpus studies to show the following: in cases where a scrambling language undergoes major changes in its clause structure, the GHC nevertheless remains an active and stable constraint in the language, indicating its status as a universal. Chapters 3 and 4 show how the GHC interacts with two independent syntactic parameters as they change over time. Chapter 3 discusses the change in the position of Tense (Tense/Infl-final vs. Tense/Infl-medial) in Yiddish (Santorini 1992, 1993) and its interaction with scrambling in that language. Chapter 4 discusses V-to-T movement, which was gradually lost in Early Modern English (Kroch 1989, Han & Kroch 2000). In both English and Yiddish, scrambling existed stably over time and is shown to exist in the modern languages, but changes in the position of the finite verb interacted with the GHC to severely restrict the potential landing sites for scrambling, even in the earliest occurrences of the innovative clause type. The results for Yiddish are simplest to describe: when Tense-medial clauses begin to appear, scrambling to clause-level (higher than Tense) is not allowed in those clauses. Additionally, clauses with objects to the left of the finite verb decline at the same rate as clauses that are unambiguously Tense-final on other criteria (slopes of -.0142 and -.0147 respectively), showing that they are the same phenomenon, and no scrambling across the verb took place. The diachronic data provides a type of information that the synchronic typology cannot: once a phrase structure change begins, the resulting variation within a single speech community and even within a single speaker immediately shows the effect of the GHC on scrambling.

In Chapter 5, I present a new antisymmetric approach to OV languages that builds on the proposals in Biberauer & Roberts (2005). The GHC is shown to help choose between a classical phrase structure and an antisymmetric one; only the latter allows the GHC to be reformulated in purely hierarchical terms, and as a general property of the adjunction operation ("Conservation of C-Command"). This chapter also shows that the Conservation of C-Command has a natural formulation in terms of Tree Adjoining Grammars (Joshi, Levy, & Takahashi 1975, Vijay-Shanker & Joshi 1985, Kroch & Joshi 1985, and much subsequent work following these original studies).

Chapter 6 continues the discussion of antisymmetry and scrambling, showing that the antisymmetric approach can cover not only the scrambling data, but also extends to other syntactic problems. In particular, the combination of an antisymmetric approach and scrambling allows a straightforward and clean account of the various patterns found in the West Germanic Verb (Projection) Raising construction, in all of its variation across the continental West Germanic dialect continuum (see Wurmbrand 2004, 2005 for an overview).

Chapter 7 discusses the Heavy NP Shift construction in English and more generally across Germanic. This construction has sometimes been taken to be a type of rightward scrambling, an analysis incompatible with the analysis of scrambling in this thesis. Additionally, HNPS appears to be so clear a case of rightward movement that it is is a challenge for any antisymmetric framework based on the ideas of Kayne (1994), such as the one presented in Chapters 5 and 6. This chapter shows that HNPS has properties that are clearly distinct from the properties of leftward scrambling and must be analyzed

differently. Even so, the analysis Kayne (1994) proposes for HNPS is also incorrect and must be rethought.

Chapter 8 addresses the issue of whether leftward scrambling to different positions and object shift really represent a unitary phenomenon. In particular, I challenge the position frequently taken that scrambling can be categorized as either A-scrambling or A'-scrambling (Mahajan 1990). I also briefly address the issue of Webelhuth's Paradox (Webelhuth 1989), and brin to light an apparently contradictory phenomenon, the Reverse Webelhuth's Paradox.

Chapter 9 is a case study from the history of English, proposing an analysis for the phenomenon of Middle English pre-Tense object clitics. Object pronouns in Middle English frequently occur to the left of the tensed verb in clauses which are otherwise VO, in apparent violation of the GHC/Conservation of C-Command. This chapter shows that these object pronouns are best analyzed as head-adjoined clitics, and this analysis would not necessarily be arrived at if it weren't for the Conservation of C-Command. Finally, Chapter 10 offers some conclusions.

Chapter 2

Functional Heads and the Typology of Scrambling

2.1 Introduction

In this chapter, I show how a broader typology of scrambling results from the interaction between the Generalized Holmberg Constraint and the inventories of functional heads that various languages have at their disposal. The hypothesis is that the GHC is a universal, but scrambling differs from language to language depending on the headedness of CP, TP, and vP. The Germanic languages have left-headed CPs, but differ from each other in the headness of TP and vP. Japanese then completes the typology, showing how scrambling operates in a right-headed CP language.

2.2 OV and VO: Comparing German and North Germanic

I begin the discussion with a comparison of the Germanic OV and the VO languages that have the simplest systems of leftward scrambling: German, Swedish, and Early Modern English. German, Swedish, and Early Modern English are all left-headed

CP languages, so the position of C is the ultimate barrier to scrambling for both the German and Swedish/EME systems. Since Swedish and EME are also left-headed TP and vP languages, the GHC also requires the verb to move to C if an element is scrambled above TP and vP. These two types of languages represent the poles of the Germanic typology, differing in all projections aside from CP. For the purposes of this discussion, I leave aside standard Dutch and the other Scandinavian languages. These languages have the same upper bound to scrambling, Comp, but they show an additional constraint as well, which is a prohibition of scrambling across the subject (Zwart 1996a, Hellan & Platzack 1995, Josefsson 1992). While the Generalized Holmberg Constraint is universal, this additional prohibition is clearly something on which languages can differ, and I take it to be related to the prohibition against scrambling over a direct object in Danish and modern American English (see discussion above). As with the direct object case, the important point here is that this additional constraint is independent of the GHC.

It is a well-known fact that weak pronouns in German can move leftward as far as the complementizer position.[2] This is illustrated by the weak pronoun *es* in sentences in (1) and (2) below for main (V2) and subordinate clauses, respectively. In both sets of sentences, the weak pronoun has left its base position (which would be to the immediate left of the verb, in a right-headed vP) to scramble leftward across some other constituent. This can be negation, an adverb, and/or the subject in a subordinate clause or a main clause with topicalization and subject-aux inversion under V2, as in (1). Note that the generalization that Comp is the barrier for pronoun scrambling holds for both main and

[2] Thanks to Tatjana Scheffler for discussion of the German data.

subordinate clauses, under the standard assumption (since den Besten 1983) that the finite verb occupies the C position in German main clauses. West Flemish is another OV-Tense-final language, which has a scrambling system that is nearly identical to standard German's (Haegeman 1996), as shown in (3) below.

(1) a. Gestern hat **es** Johann nicht gekauft.
Yesterday has it Johann not bought.
"John didn't buy it yesterday."

b. Gestern hat Johann **es** nicht gekauft.
Yesterday has Johann it not bought.

(2) a. Ich weiß daß **es** Johann gestern nicht gekauft hat.
I know that it Johann yesterday not bought has.
"I know that John didn't buy it yesterday."

b. Ich weiß daß Johann **es** gestern nicht gekauft hat.
I know that Johann it yesterday not bought has.

Like German, weak object pronouns in West Flemish move leftward from their base positions to a number of landing sites, with the leftmost possible landing site being to the right of C:

(3) a. ... da **ze** Valère Marie misschien gegeven eet.
that them Valère Mary perhaps given has.
"... that Valère has perhaps given them to Mary."

b. ... da Valère **ze** Marie misschien gegeven eet.
that Valère them Mary perhaps given has.

c. ... da Valère Marie **ze** misschien gegeven eet.
that Valère Mary them perhaps given has.
(Haegeman 1996: 150)

In fact, for many speakers of German, the leftmost position adjacent to the complementizer is the most natural position (or nearly obligatory) for weak object pronouns. This was also apparently true of the early Old English of Beowulf, which had a system of scrambling very similar to that of modern German or West Flemish at a time when a sizeable proportion of Old English clauses were still OV and Tense-final (Pintzuk 1996). This is an important fact which I will return to in Chapter 4 especially, where I discuss the fact that English eventually changed from the earlier scrambling system into a VO object shift language of the modern Swedish type.

Scrambling to any position to the left of C, on the other hand, is entirely impossible for German speakers. Thus, object pronouns can never appear to the left of a complementizer or a finite verb:

(4) a. * **Es** gestern hat Johann nicht gekauft.
It yesterday has Johann not bought.

b. * Gestern **es** hat Johann nicht gekauft.
Yesterday it has Johann not bought.

(5) * Ich weiß **es** daß Johann gestern nicht gekauft hat.
I know it that Johann yesterday not bought has.

For the purposes of this discussion, the important point is that weak pronouns must scramble leftward from their base positions (just as I argued for English weak pronouns above) and that in German, this leftward scrambling is bounded by the merger of a head

in C.[3] The observation I would like to emphasize here is that the GHC is not about the position of verbs per se, which is how it is usually viewed in the Scandinavian object shift literature (a notable exception is Holmberg 1999). It is about the position of functional heads, and more specifically, about functional heads into which something has been merged. Thus, in German, the barrier to scrambling is either a complementizer or a finite verb (whether a lexical verb or an auxiliary, as both move to C in German).

In VO languages, on the other hand, the barrier is usually marked by the position of the verb, though I will discuss a few cases in the full dissertation where the barrier is due to the merger of some other item with a functional head. And of course, because nonfinite verbs also appear to the left of objects in VO languages, pronouns can be trapped in much lower positions than they are in German; present-day English is an extreme example of this, as I discussed above. However, when the verb has moved as far as possible (i.e. Comp), pronouns in a Swedish or EME-style system can potentially scramble to as high a structural position as their German or West Flemish counterparts do. It is also interesting to note that the statistics for Early Modern English and Middle English suggest that long-object-shift to the C-adjacent position was not only possible, but highly preferred to leaving the object in some lower position when the verb has

[3] At the moment, I am only concerned with showing where the leftmost landing site is. The question as to why some lower potential vP-adjoined landing sites are not possible, and why speakers/dialects differ as to which lower landing sites are available, I leave for later research. This is actually a question that is independent of the key parameter of variation I'm trying to highlight here, namely, the difference between the OV and VO systems of scrambling, as these speaker/dialect difference are also found among Scandinavian speakers in the VO object shift system. I will note here, however, that the stipulation that only some landing sites are possible for scrambling, and the variation in these landing sites from language variety to variety, are both reminiscent of the restrictions on adjunction sites for adverbs and the variation in adverb placement from variety to variety. As I will argue below in chapter 5, this and many other characteristics of scrambling are best understood as general characteristics of adjunction, with scrambling being a type of adjunction (following Lee 1993 and others).

moved to C, just as in German. A Swedish example is given in (6) below and an EME example is in (7) (I also show a number of other examples of the long-object-shift type in Chapter 4 below for Early Modern English and modern Swedish).

(6) Klarar **sig** barnen på egen hand?
manage REFL children-the on own hand
"Do the children manage on their own?"
(Hellan & Platzack 1995: 58)

(7) And thither bringes **you** a naturall instinct to true goodnes,
"And a natural instinct brings you there to true goodness"
(Queen Elizabeth I's English translation of Boethius' *De Consolatione Philosophiae*, BOETHEL-E2-P2,48.73 in PPCEME, date: 1593)

Also, like German or West Flemish, Swedish allows weak pronouns to surface in a number of lower landing sites as well, yielding the full paradigm in (8a-e).

(8) a. Igår läste han **dem** ju alltså troligen inte.
Yesterday read he them as-you-know thus probably not.

b. Igår läste han ju **dem** alltså troligen inte.
c. Igår läste han ju alltså **dem** troligen inte.
d. Igår läste han ju alltså troligen **dem** inte.
e. Igår läste han ju alltså troligen inte **dem**.
(Hellan & Platzack 1995)

From (8e) we can see that Christer Platzack's dialect/idiolect of Swedish even allows weak object pronouns to remain in a landing site below negation (though there is some variation between Swedish speakers on this point – Christer Platzack p.c.). I take the object to be adjoined to RootP in this type of example, which is presumably the

lowest landing site available that is still to the left of the object's base position. This fact shows Swedish to be very much like present-day English, as well as Early Modern English (see the discussion of modern English in section 4.2). The large number of landing sites shown in the Swedish data in (8) also indicates that scrambled/object-shifted pronouns should be analyzed as adjoining to phrasal projections, under the standard assumption that *ålltså*, *troligen*, and the negation *inte* in (8) are vP-adjoined adverbs. Otherwise, in a system where scrambling/object shift lands objects in the specifiers of functional projections, for instance, it would be mysterious why there are just as many landing sites for the object as their are adverbs. (This argument actually applies equally to the OV and VO languages, as well.)

2.3 An Intermediate System: Yiddish and Kru

If German and Swedish represent the poles of the scrambling typology, showing generalized scrambling systems as they are affected by an OV-Tense-final and a VO-Tense-medial clause structure, then modern Yiddish represents an important intermediate case. It is well-established that Yiddish changed during its history from a Tense-final language to become a left-headed-Tense language (Santorini 1992, Santorini 1993a): tensed verbs in modern Yiddish always appear to the left of their complements in both main and subordinate clauses. The structure of the Yiddish vP, on the other hand, is by no means a settled issue, since certain DP objects and nonfinite auxiliaries may be found

to both the left and the right of nonfinite lexical verbs in the modern language. There is a long-standing debate in the literature as to whether modern Yiddish should be analyzed as underlyingly VO, with other orders being derived by leftward movement processes (e.g., as in Diesing 1997), underlyingly OV, with other orders derived by rightward extraposition processes (e.g. as in Hall 1979, Geilfuß 1991, Vikner 2001, and references therein), or a language containing both OV and VO VPs (as in Santorini 1993b). It is beyond the scope of this proposal to discuss all of the details of this debate (cf. Santorini 1993b for a brief summary), but I argue, along the lines of Vikner (2001), that the preponderance of evidence is on the side of Yiddish being an OV language.

The statement of the Generalized Holmberg Constraint, however, actually forces our hand: it has as its consequence that Yiddish is OV, a position that has considerable independent empirical support anyway (as shown in the references above). Preverbal DPs in Yiddish can be shown to scramble leftward over negation and adverbs, as in the following examples.[4]

(9) a. Ikh trakht az Hayim hot **dem bikhl** nekhtn nit gekoyft.
I think that Hayim has the book-DIM yesterday not bought
"I think that Hayim didn't buy the book yesterday."

b. Ikh trakht az Hayim hot nekhtn **dem bikhl** nit gekoyft.
I think that Hayim has yesterday the book-DIM not bought.

This is predicted by the Generalized Holmberg Constraint, but only if the DPs are base-generated preverbally. Scrambling across a verb (i.e., across a verb Root which has been

[4] I would like to thank Abraham Zeif for this example, as well as for a number of other Yiddish judgments and examples below. Mr. Zeif is a native speaker of a variety of Lithuanian Yiddish, from the shtetl of Jody, Poland. All Yiddish examples and judgments in this section that are not otherwise attributed are Mr. Zeif's.

merged in a functional head, little-*v*) is predicted to not exist under the Holmberg Constraint, and this prediction is borne out in the rest of Germanic. The situation is clearest with weak pronouns in Yiddish, which must scramble leftward, and cannot occur after a nonfinite verb in Yiddish; cf. the contrast below between the grammatical (9a,b) and the ungrammatical (9c). And note that as in German and Swedish, the positions in which the weak pronoun *im* may be found are vP-adjoined.

(10) (Context: Hot Hayim nekhtn gekoyft a bikhl?
Has Hayim yesterday bought a book-DIM?)

a. Ikh trakht az Hayim hot **im** nekhtn nit gekoyft.
I think that Hayim has him yesterday not bought.
"I think that Hayim didn't buy it yesterday."

b. Ikh trakht az Hayim hot nekhtn **im** nit gekoyft.

c. * Ikh trakht az Hayim hot nekhtn nit gekoyft **im**.

Again, this is characteristic of Germanic OV languages with rightward extraposition and/or Heavy NP Shift: weak pronouns, unlike full DPs, are prosodically too light to shift rightward across the verb, and so they are only found in some preverbal position. The syntax of weak pronouns in Yiddish has no natural explanation under the hypothesis that the Yiddish vP is VO, because all of the Germanic VO languages show postverbal weak pronouns (whether the verb is finite or nonfinite). If on the other hand, we suppose that modern Yiddish has a left-headed TP but a right-headed vP, then the Generalized Holmberg Constraint predicts precisely this distribution of weak pronouns.

We can see that the Generalized Holmberg Constraint operates in Yiddish independently of the structure of the vP, because modern Yiddish is Tense-medial, rather

than Tense-final like German (as shown in Santorini 1992, 1993a). For this reason, weak pronouns in modern Yiddish can only move leftward as far as the position of the finite verb in Tense, not all the way to C as in German or West Flemish. The subordinate clause examples in (10) above and (11) below show this unambiguously; C is filled by the complementizer, so the finite auxiliary must be in Tense.

(11) a. * Ikh trakht az Hayim **im** hot nekhtn nit gekoyft.
I think that Hayim him has yesterday not bought

b. * Ikh trakht az **im** Hayim hot nekhtn nit gekoyft.
I think that him Hayim has yesterday not bought

The positions to the left of the auxiliary *hot* ("has") in the examples above are unavailable as landing sites for scrambling, because movement to those positions would entail crossing a functional head in which an element has been merged, in violation of the GHC. This is in stark contrast to the German subordinate clause examples in (2) above, in which the most natural position for weak pronouns is immediately adjacent to C. This position is available in a German subordinate clause precisely because German is Tense-final, and so C is the only functional head to the left of the object's base position.

The analysis of Yiddish (left-headed TP, but right-headed vP) does not make it a typological loner, either, though it is unusual within Germanic. Rather, it patterns with other "SIOV" languages found elsewhere in the world, such as the Kru languages Vata and Gbadi spoken in Ivory Coast (described in Koopman 1984). These languages have a left-headed TP with obligatory movement of the finite verb to Tense, but a right-headed vP, which is visible in the position of nonfinite verbs. Interestingly, Vata and Gbadi also

show leftward scrambling of DPs. As predicted by the Generalized Holmberg Constraint, this leftward scrambling of DPs is bounded by the position of the finite verb or auxiliary in Tense (according to the description in Koopman 1984: 27-29). The scrambling possibilities in Kru are shown in the Vata sentences below, taken from Koopman (1984: 29).[5]

(12) a. n^3 ka^4 yO^3-O^3 sle^4-e^3 mlI^4 sa^3ka^4 nyE^3
I FUT-A child-DEF house-DEF in rice give
"I will give rice to the child in the house."

b. n^3 ka^4 sle^4-e^3 mlI^4 yO^3-O^3 sa^3ka^4 nyE^3
I FUT-A house-DEF in child-DEF rice give

c. n^3 ka^4 sle^4-e^3 mlI^4 sa^3ka^4 yO^3-O^3 nyE^3
I FUT-A house-DEF in rice child-DEF give

d. n^3 ka^4 sa^3ka^4 sle^4-e^3 mlI^4 yO^3-O^3 nyE^3
I FUT-A rice house-DEF in child-DEF give

e. n^3 ka^4 yO^3-O^3 sa^3ka^4 sle^4-e^3 mlI^4 nyE^3
I FUT-A child-DEF rice house-DEF in give
"I will give rice to the child in the house."

f. n^3 ka^4 sa^3ka^4 yO^3-O^3 sle^4-e^3 mlI^4 nyE^3
I FUT-A rice child-DEF house-DEF in give

Although Kru differs from Germanic in not having extraposition of direct object DPs (Heavy NP Shift; see discussion in Chapter 7), Kru does show extraposition of finite

[5] For convenience, I transcribe the Vata tones as superscripts 1-4, corresponding to low, mid, mid-high, and high, respectively. Otherwise, I follow the transcription system in Koopman (1984)

clause complements and PPs across a nonfinite verb in its base OV position, as shown in (13) and (14) below, respectively.[6]

(13) n^{3} nI^{4} gu^{1}gu^{1} **na^{2} wa^{3} nI^{3} yi^{2}**
I NEG-A believe NA they FUT-A come
"I did not believe that they were coming."
(Koopman 1984: 109)

(14) a. a^{1} nI^{2}-ka^{2} yue^{4} sa^{3}ka^{4} **sle^{4}-e^{3}** **mlI^{4}** nyE^{3}
we FUT-A FT children rice house-DEF in give

b. a^{1} nI^{2}-ka^{2} yue^{4} sa^{3}ka^{4} nyE^{3} **sle^{4}-e^{3}** **mlI^{4}**
we FUT-A FT children rice give house-DEF in

Aside from the lack of DP extraposition/HNPS, the Kru system looks strikingly similar to the Yiddish system: it is OV, Tense-medial, extraposes heavy phrases rightward across nonfinite verbs, has leftward scrambling bounded by the position of the finite verb in Tense, and shows no evidence of leftward scrambling across verbs. In this way, scrambling in both Yiddish and Kru is restricted to a lower structural position than it is in German by the addition of one more left-headed functional projection.

2.4 Continuing the Argument that Yiddish is OV

[6] Note that I will continue to refer to extraposition constructions, including Heavy NP Shift, as "rightward movement" throughout the first four chapters of this dissertation. However, in Chapter 7, I propose a new analysis of Heavy NP Shift as leftward movement, in accordance with the antisymmetric approach to phrase structure adopted in Chapter 5. An antisymmetry-compatible analysis of the West Germanic Verb (Projection) Raising construction is presented in Chapter 6.

The OV analysis of Yiddish falls entirely in line with the predictions of the GHC and the typology of scrambling languages laid out in this dissertation. In addition to this indirect evidence, the analysis in Hall (1979) and Geilfuß (1991) seems is the Occam's Razor hypothesis for modern Yiddish independently of the GHC: it covers the data while creating the fewest new theoretical entities. Geilfuß shows that it is possible to derive all of the modern Yiddish word orders if one assumes a uniformly OV VP, along with a set rightward movement processes which are independently known to exist in Germanic: Heavy NP Shift, PP extraposition, and the West Germanic verb-raising construction (on the last, cf. Evers 1975, Zaenen 1979, Haegeman & van Riemsdijk 1986, Kroch & Santorini 1991, Wurmbrand 2004, 2005 and references therein, *inter alia*). These rightward movement operations are known to produce what appear to be VO orders on the surface in modern Germanic languages that are well-established to be underlyingly OV on other grounds (e.g. Dutch and vernacular regional varieties of German). Furthermore, each one of these rightward-movement processes is unambiguously attested in earlier stages of Yiddish (Santorini 1992, 1993a). This last piece of information was not one that Geilfuß or Hall (1979) necessarily had access to, but it makes the appeal to rightward-movement in modern Yiddish all the more plausible. As I mentioned above, Santorini (1992, 1993a) showed the Yiddish changed from a Tense-final (or Infl-final) language, like modern Dutch or German, to a left-headed-Tense (or Infl-medial) language. During the course of this study, Santorini observed that

earlier Yiddish showed unambiguously Tense-final clauses[7] which also showed the application of one or more of the rightward movement processes mentioned above. Because these sentences were unambiguously derived from an underlying Tense-final clause structure, they clearly show that early Yiddish allowed the movement of some nonfinite verbs (verb-raising), PPs, and DP objects rightward across a verb.[8] Examples of these three, respectively, are shown below: [9]

(11) dr veyl es gimeyniklikh iz **giv[o]rdn**
because it common is become
(Santorini 1992: 607; example from Anshel ben Joseph's Preface to *Merkevet ha-mishneh*, 1r, date: 1534)

(12) d[a]z ikh reyn verde **fun der ashin**
that I clean become from the ashes
(Santorini 1992: 607; example from Johann Jakob Christian's *Eyn sheyn purim shpil*,1004, date: 1697)

(13) ven er nit veys **eyn guti veyd**
if he not knows a good pasture
(Santorini 1992: 607; example from Abraham Apotheker Ashkenazi's *Sam hayyim*, 41, date: 1590)

Santorini (1993a) also showed that the rate of extraposition for DP objects did not change significantly over time during the course of its history, and so it is plausible that it still applies in the modern language, even though modern Yiddish is no longer Tense-final (cf. also Kroch & Pintzuk (1989) for evidence that modern English Heavy NP Shift is the

[7] Underlying Tense-final structure was diagnosed by the presence of the subject and some other element (e.g. an object or negation) preceeding the finite verb, or verb-aux order in clauses containing auxiliaries which did not exhibit verb-raising. See also Pintzuk (1991).

[8] See footnote #6 above.

[9] For examples of historical Yiddish, I use the orthography/romanizations given in Santorini (1992, 1993a). For modern examples, I have tried to keep a consistent orthography along the lines of the YIVO romanization.

same movement process as Heavy NP Shift in early Old English, which also applied across verbs in Tense-final position).

Thus, if one maintains an analysis of modern Yiddish VPs as right-headed rather than left, it is possible to derive modern Yiddish sentences like the ones below by only appealing to movement processes that are both attested across Germanic and attested in the history of Yiddish itself.

(14) Ikh hob gezen **Moishn**
I have seen Moses

(15) Er hot geleynt **dos bukh**
He has read the book.
(examples from den Besten & Moed-van Walraven 1986: 125)

On the other hand, if one treats Yiddish as underlyingly VO, the only way to derive examples like (16) and (17) below is by a rule that generally scrambles objects leftward across the nonfinite verb (this is the approach in Diesing 1997). According to the Generalized Holmberg Constraint in (10), this type of scrambling should not exist, and there is a very good reason to believe that the Constraint is making the correct prediction here: leftward scrambling across a verb is unattested (at least) in Germanic.[10]

[10] This is true of non-negative, non-quantified DPs. There is a leftward movement process which moves only negated and quantified objects across a nonfinite verb that is attested in Late Middle English (van der Wurff 1998, Kroch & Taylor 2000), which was almost uniformly VO in the relevant time period, and in modern Icelandic (Rögnvaldsson 1987), which is underlyingly VO under every diagnostic. I take this process to be a different type of movement from the general DP and pronoun scrambling discussed here, and in fact, Jónsson (1996) showed that it is a type of A'-movement in modern Icelandic, and Light & Wallenberg (2008) showed that it is a type of A'-movement to Spec(MoodP) in a Split-Infl phrase structure in both Icelandic and earlier stages of English.

(16) Ikh hob **Moishn** gezen
I have Moses seen

(17) Er hot **dos bukh** geleynt
He has the book read
(examples from den Besten & Moed-van Walraven 1986: 125)

As we have already seen above, scrambling of pronouns (and, indeed, DPs) in both OV and VO languages obey the Generalized Holmberg Constraint and do not cross (filled) heads to their left. Icelandic is a particularly good example, as it has scrambling of definite DPs, like the DPs in the Yiddish examples above, but neither DPs nor pronouns can cross a verb to their left. The mainland Scandinavian languages do not have DP scrambling at all, but the elements they can scramble (pronouns) obey the Generalized Holmberg Constraint, as shown for Swedish above. There is also good evidence from Old English and Early Middle English that when DPs scramble, they nevertheless do not scramble from an underlying postverbal position to a preverbal surface position. I will return to the facts of Old English and Middle English in great detail in the next few sections of this proposal, but this fact is important enough to the discussion at hand that I will point it out in advance of the rest of my discussion of English. Old English and Early Middle English were in transition from an OV to a VO grammar, and so it is possible to find both underlyingly OV and VO vP/VPs in these stages of the language (cf. Pintzuk 1991, Kroch & Taylor 2000, *inter alia*). However, underlyingly left-headed VPs may be found by looking for a number of diagnostic elements, light elements that are almost never extraposed in OV Germanic languages (e.g. verbal particles, weak pronouns). Pintzuk (2002) included a study of all of the clauses with such diagnostic

elements and nonfinite verbs in the York-Toronto-Helsinki Parsed Corpus of Old English ("YCOE"; Taylor, Warner, Pinztuk & Beths 2002). She found that out of 32 clauses containing a postverbal (VO) diagnostic element and some nominal, non-quantified[11] DP object in YCOE, there were zero clauses in which the nominal DP preceded the nonfinite verb. In other words, where it's possible to tell that a clause is VO, no DPs occur to the left of the verb. Kroch & Taylor (2000) report the same result for Middle English, with 0 (out of 19) tokens in their sample showing a DP to the left of the nonfinite verb where there is also a VO diagnostic.

In addition to the fact that Diesing's (1997) analysis of Yiddish as VO must appeal to a type of scrambling for which there is no evidence, data from German shows her primary argument to be inconclusive. Diesing argues on semantic grounds that preverbal DPs in Yiddish cannot be *in situ* inside the VP. She notes that there is a strong dispreference for definite DPs to remain *in situ* inside the VP in German and Dutch, and that they generally scramble leftward in those languages if the DP is not contrastively stressed and there is some other element in the clause that they can scramble past (e.g. negation or adverbs). This is not the case with indefinite DPs, which can more frequently remain inside the VP and to the right of sentential adverbs and negation in the Germanic OV languages. Diesing found that, according to her speakers, this same contrast holds in Yiddish between preverbal (pre-nonfinite-verb) and postverbal position in clauses containing auxiliaries and nonfinite verbs: her informants prefer definite DPs in preverbal position and indefinite DPs in postverbal position, and if sentenial adverbs are present,

[11] See last footnote.

definite DPs occur to the left of those as well. From this set of facts, she concludes that the preverbal DPs are scrambled from an underlying VO position, and the postverbal DPs are *in situ* in the VP.

First of all, let me point out that even under an analysis that Yiddish is uniformly OV, these preverbal DPs need not be analyzed as being *in situ*. Just based on the formal properties of the syntactic framework that Diesing and I are both assuming, if a DP can scramble leftward at all, there is no reason it couldn't scramble leftward string-vacuously. And regardless of how one analyzes the Yiddish vP/VP, one must admit the fact that Yiddish definite DPs can scramble leftward past other clausal constituents, such as sentential negation and vP-adjoined adverbs, as shown by the position of *the book* in the sentences below.[12]

(18) a. Ikh trakht az Hayim hot **dem bikhl** nekhtn nit gekoyft.
I think that Hayim has the book-DIM yesterday not bought
"I think that Hayim didn't buy the book yesterday."

b. Ikh trakht az Hayim hot nekhtn **dem bikhl** nit gekoyft.
I think that Hayim has yesterday the book-DIM not bought.

If the Yiddish vP/VP is in fact right-headed rather than left-headed, then examples with preverbal definite DPs such as (16), (17), and (18) are simply parallel to German examples such as those below.

[12] I would like to thank Abraham Zeif for this example, as well as for a number of other Yiddish judgments and examples below. Mr. Zeif is a native speaker of a variety of Lithuanian Yiddish, from the shtetl of Jody, Poland. All Yiddish examples and judgments in this section that are not otherwise attributed Mr. Zeif's.

(19) Er hat Johann gesehen.
He has John seen

(20) Ich habe das Buch gelesen
I have the book read

Since Diesing agrees that German's VP is underlyingly OV, and also maintains that definite DPs must scramble leftward out of the VP, she would have to analyze this type of German example as involving string-vacuous scrambling of the DP, as they are certainly not ineffable for German speakers.

Rather than assuming that the indefinite DPs are *in situ*, I would take the definiteness effect that Diesing found to indicate that indefinite DPs are more likely to extrapose rightward (i.e. Heavy NP Shift) than definite DPs are, and that Yiddish is an OV language with a great deal of rightward movement of DPs.[13] This is not necessarily a stipulation about Heavy NP Shift, either. I am glad to accept Diesing's conclusion that definite DPs prefer to scramble leftward, as Yiddish undeniably does have DP-scrambling (as in 18), but I would like to suggest that leftward scrambling simply bleeds rightward extraposition. It seems likely that speakers would not (or could not) scramble and extrapose a DP simultaneously, particularly since the phonological requirements on the two operations, such as stress and heaviness, are well-known to be roughly opposite to each other. The definite DPs in sentences like (16) and (17) I would analyze as having scrambled string-vacuously from an underlying OV position, parallel to the German examples in (19) and (20), and this predicts a pattern in which definite DPs are

[13] Again, see footnote #6 above.

extraposed to the right of the verb less often than indefinite DPs. Indefinite DPs, on the other hand, do not need to scramble, and so they may be freely extraposed. This accounts for Diesing's observation that postverbal DPs in Yiddish are more likely to be indefinite than definite. I will also note here, however, that it is important to consider the pragmatic context when analyzing the facts of DP placement in Yiddish. It is not the case that definite DPs in Yiddish cannot be found postverbally; there is no hard requirement that they scramble leftward, but rather that they frequently undergo scrambling (and escape Heavy NP Shift) given a certain context. Abraham Zeif (p.c.; see last footnote), for instance, judges the two sentences below as equally acceptable, but said that he would only use the (21) (with the preverbal DP) as an answer to a question about *the book*, or in a discourse where the entity *the book* had been previously mentioned.[14] The version with the postverbal object in (22), on the other hand, carries no such implication.

(21) Ikh hob dem bikhl gekoyft.
I have the book bought.

(22) Ikh hob gekoyft dem bikhl.
I have bought the book.

By the same token, the pragmatic pressure to scramble given information leftward will disfavor preverbal indefinite DPs in a language with relatively free rightward extraposition of un-scrambled elements, as indefinite DPs generally introduce new

[14] As I recall, Abe's exact words to me were, "...ober dos is an entfer af a frage. Velkhe bikhl?! Vos redn mir vegn?!" ("...but this is an answer to a question. Which book?! What are we talking about?!")

entities into a discourse. This also seems to be true of German varieties that allow greater freedom of rightward extraposition than standard German does, though they are unambiguously OV and Tense-final (which can be seen by, e.g., the final position of the finite verb in subordinate clauses). Augustin Speyer (p.c.) reports that for his native dialect, in colloquial speech, DP objects may be moved to the right of a nonfinite verb, but that indefinite objects sound somewhat more natural in the postverbal position than definite objects do. However, as I am suggesting, this is not a hard constraint of the syntax, but rather a pragmatic preference. Augustin Speyer (p.c.), for instance, considers both of the following sentences to be grammatical even though the DPs differ in definiteness, just as Abraham Zeif's Yiddish also allows postverbal definite and indefinite DPs. The objects do have to be stressed and narrowly focused, however (see Chapter 7).

(23) Ich hab dann gestern auf den Markt geschleppt meine Hennen
I have then yesterday to the market dragged my hens
"I dragged my hen to market yesterday."

(24) Ich hab dann gestern auf den Markt geschleppt eine Henne
I have then yesterday to the market dragged a hen
"I dragged a hen to market yesterday."

When nothing occupies the "Mittelfeld" between the auxiliary and nonfinite verb, as in the Yiddish examples in (14), (15), and (22) above, Heavy NP Shift in colloquial German becomes degraded for the speakers I have consulted. However, it seems to be more degraded if the shifted DP is definite:

(25) ? Ich habe gelesen ein Buch
I have read a book

(26) ?* Ich habe gelesen das Buch
I have read the book

Given that this variety is underlyingly OV, any definiteness effect (to the extent that it really exists) cannot be attributed to some fact about the DP object's base position, but rather must be seen as an effect of how Heavy NP Shift interacts with other aspects of the grammar that are sensitive to definiteness, e.g. scrambling. This is the approach I choose to take for Yiddish as well.

Finally, a brief note on the ordering of Yiddish nonfinite verbs. As Diesing (1997) and others note, Yiddish shows both typical OV and VO patterns in the ordering of nonfinite verbs in clauses containing two or more nonfinite verbs. The two patterns are shown in (27) and (28) below (again, these examples are due to Abraham Zeif, and these occurred in natural, running speech).

(27) Vilstu geyn shpatziren?
Want-you go walking
"Do you want to go for a walk?"

(28) Er iz geboren gevoren in de milkhome
He is born became in the war
"He was born in war [i.e., war is second nature to him]"

(29) Der Kind iz farfalen gegangen.
The child is lost gone.
"The child went missing."

(27) shows the same ordering of nonfinite verbs found in VO languages, such as English or Swedish, while (28) has the nonfinite passive auxiliary *gevoren* following the lexical past participle *geboren*, as it would be in standard German. Diesing (1997) analyzes (27) as the base order and (28) as a derived order. I would simply remark, however, that regardless of how these sentences are derived, OV auxiliary orders such as the one found in (28) are entirely unattested in the Germanic VO languages, i.e. English and the Scandinavian languages.

As in the case of preverbal DP objects, whatever the details of the operation are that would derive (28) from a VO base, Diesing has to propose some process that is not found outside of Yiddish in order to account for the order of verbs in (28). The order in (27), on the other hand, is not unattested in the Germanic OV languages. On the contrary, it is the order of nonfinite verbs that is characteristic of the West Germanic verb-raising construction, found in Old English (Pintzuk 1991), Dutch, and Swiss/South German dialects (Kroch & Santorini 1991, Zaenen 1979, among others; see section 6.3 below), and older stages of Yiddish (Santorini 1992, 1993a). Thus, appealing to verb-raising (analyzed in the above references as extraposition of a nonfinite verb) in order to derive (27) from an OV base for Yiddish, is only making use of an operation that is already necessary in order to derive a number of constructions in well-established OV languages. (28) would then represent the base order, the normal order for OV auxiliaries, as found in standard German. Furthermore, Dutch shows that verb-raising is lexically governed, and obligatory with certain verbs; a case in point is nonfinite clausal complements, such as (27). It is therefore not too surprising that verb-raising should be

obligatory in Yiddish in a number of contexts, nor is it surprising that a particular auxiliary, the passive auxiliary *vern* (*gevoren*) should be lexically marked to resist it.

In addition to filling out the Germanic typology, Yiddish also demonstrates the Generalized Holmberg Constraint in a dynamic way, a topic I take up in detail in the next chapter. I will simply point out here that Yiddish changed over time from a German-style Tense-final clause structure to a Tense-medial clause structure, and as it did so, scrambled DPs became trapped in lower structural positions due to the presence of the finite verb in Tense. As the change in Tense went to completion, objects could no longer move as far as C, as they did in older stages of Yiddish, unless the finite verb also moved to C. In the next chapter, I examine this change in detail and compare it to a similar change observed by Taylor (1990) in Ancient Greek.

2.5 Complementizer Final Languages Allow Long Scrambling

One final piece of the typological jigsaw is provided by looking at scrambling outside of Indo-European, in more relentlessly head-final languages such as Japanese. Of the Germanic languages, German and West Flemish show the most free type of scrambling, in the sense that their left-moved elements surface in the highest structural positions of all the languages in both root and embedded contexts, and whether or not the lexical verb is finite. As I discussed above, this behavior exists because the Germanic OV languages contain only a single left-headed functional projection, CP, and so C (with

a verb or complementizer merged) is the only barrier to leftward movement. This analysis, which I proposed initially based on the Germanic data alone, makes the further prediction that if a language lacked a left-headed CP, and complementizers were merged in a right-headed functional projection instead, then CP would not be a barrier to scrambling. This means that if such a language allows leftward scrambling generally, then scrambled phrases should be able to move leftward out of finite embedded clauses with overt complementizers.

This is precisely the situation found in Japanese, which contains complementizer-final CPs rather than left-headed CPs (i.e. all clause-level functional projections are right-headed in Japanese). As the examples show below, it is possible not only to generally scramble in Japanese, but also to scramble DPs and PPs that originate in a finite complement clause leftward past constituents of the matrix clause (Saito 1992, Saito & Fukui 1998, and references therein). (30a) shows the base order, with constituents of the embedded clause in situ, and (30b,c) show two possible orders derived by scrambling out of the embedded clause (scrambled phrases are in boldface). Note the clause-final position of the complementizer, *to*.

(30) a. Bill-ga Mary-ga John-ni sono hon-o watasita to itta (koto)
Bill-NOM Mary-NOM John-to that book-ACC handed that said (fact)
"Bill said that Mary handed that book to John."

b. **Sono hon-o** **John-ni** Bill-ga Mary-ga watasita to itta (koto)
that book-ACC John-to Bill-NOM Mary-NOM handed that said (fact)

c. **John-ni sono hon-o** Bill-ga Mary-ga watasita to itta (koto)
John-to that book-ACC Bill-NOM Mary-NOM handed that said (fact)
(Saito & Fukui 1998: 443-444)

Japanese scrambling can be shown to be a general leftward movement process rather than a usual case of A' movement to some particular Spec position by the fact that there is no upper bound on the number of phrases that can be scrambled at the same time. This is shown in the examples below:

(31) Soko-de$_k$ John-ni$_j$ sono hon-o$_i$ Bill-ga Mary-ga t_k t_j t_i watasita
there-at John-to that book-ACC Bill-NOM Mary-NOM handed

to sinziteiru (koto)
that believes (fact)
"Bill believes that Mary handed that book to John there."

(32) Mikka-mae-ni soko-de$_k$ John-ni$_j$ sono hon-o$_i$ Bill-ga
three-days-before-at there-at John-to that book-ACC Bill-NOM

Mary-ga t_l t_k t_j t_i watasita to sinziteiru (koto)
Mary-NOM handed that believes (fact)
"Bill believes that Mary handed that book to John there three days ago."
(Saito & Fukui 1998: 444, footnote #8)

While a full discussion of so-called "long scrambling" in Japanese and Korean is beyond the scope of this dissertation (see Saito 1985, Saito 1992, Saito & Fukui 1998, and Lee 1993 and references therein for the same phenomenon in Korean), I will observe that this phenomenon is entirely expected under the GHC; it does not require further theoretical machinery to be accommodated. It is simply a consequence of the same scrambling operation at work in Germanic, but placed in the context of a complementizer phase structure.

2.6 Summary

With the addition of Yiddish and Kru as OV-Tense-medial and Japanese and Korean as Comp-final, we now have a complete typology of scrambling systems, defined by three parameters: the position of Comp, Tense, little-*v*, and how far finite verbs move. In all of the cases so far, the scrambling rule itself is simple: move left to left-adjoin to a phrasal projection without violating the Generalized Holmberg Constraint. Icelandic is VO, but with V-to-T movement and V-to-C movement, and so scrambling occurs in both main and subordinate clauses, bounded by T in subordinate clauses, and either T or C in main clauses. English, on the other hand, as I show later on in Chapter 4, demonstrates over the course of its history how a change in one parameter, verb movement to T, interacts with the Generalized Holmberg Constraint to affect the scrambling possiblities of a language. Modern (American) English is VO without either V-to-C movement or V-to-T movement (of lexical verbs). Thus all objects, including weak pronouns, are always trapped in a low position. Belfast English (as we will see below) is like American English, but with V-to-C movement in only one context, imperatives, and so it shows scrambling in only that context (Henry 1995). Swedish (like most mainland Scandinavian languages) is VO with V-to-C movement in main clauses and no V-to-T movement, and so it shows scrambling just in main clauses, with the set of landing sites for scrambling bounded by C (Hellan & Platzack 1995: 52).

Chapter 3

Scrambling and Phrase Structure Change in Yiddish

3.1 Introduction: Consequences for scrambling of the Tense-final to Tense-medial change

In addition to filling out the Germanic typology, Yiddish also demonstrates the Generalized Holmberg Constraint in a dynamic way. As Santorini (1992, 1993a) showed, Yiddish gradually changed from a German-like Tense-final language into a Tense-medial language (or left-headed TP language, under classical X-bar theoretic assumptions) roughly during the 15th-18th centuries. After the change was initiated, and Tense-medial TPs were introduced into the Yiddish speech community, a period of variation began in which there were Tense-final and Tense-medial TPs produced both by the community and by individual speakers (see Santorini 1992, where this fact is established beyond doubt). In the study below, I show in two different ways that as soon as Tense-medial clauses begin to appear, scrambling to clause-level (higher than Tense) is not allowed in those clauses. Thus, it is clear that a ban on scrambling past Tense did not need to evolve in Yiddish in coordination with the Tense-final to Tense-medial

change as an additional change. Rather, scrambling past Tense in a Tense-medial clause is due to the GHC, and the GHC is such an integral part of UG that the moment a Tense-medial clause is formed in the mind of a speaker, it is an obvious structural necessity that scrambling be constrained by the Tense head on the left.

3.2 Scrambling in Early Yiddish

Under its original German-like grammar, Yiddish showed Tense-final clauses with objects frequently scrambled to a high, TP-adjoined position in addition to lower landing sites for scrambling. In the examples of Tense-final subordinate clauses below, the object in boldface scrambles to a landing site above the subject and adjacent to C, just as in modern German. Note also that these clauses are unambiguously Tense-final, which can be determined by the presence of some diagnostic element in the sentence. As Santorini (1992, 1993a) discusses, many of the clauses in the Yiddish historical data during the period when the position of Tense was changing are ambiguous and could be surface representations of either underlyingly Tense-medial or Tense-final clauses. In (2) and (3) the position of the particles *moykhl* and *oyz*, preceding the tensed verb, is diagnostic of a Tense-final clause (certain elements borrowed from Hebrew behave as Germanic particles in Yiddish; see Santorini 1992). (1) and (3) both contain the Tense-final diagnostic of a nonfinite verb preceding the finite auxiliary, and in (1), there is the additional Tense-final diagnostic of the negation preceding the finite verb: the auxiliary

has not moved to the left of negation by V-to-T movement to medial Tense. (See also Pintzuk 1991, Kroch & Taylor 2000, Pintzuk 2005, Pintzuk & Taylor 2004, Pintzuk & Haeberli 2006, for the use of diagnostic elements in analyzing Old and Middle English.)

(1) ... d[a]z **mir** yusf di h' zhubim nit gebn vil
that me Joseph the five guilders not give wants
"that Joseph doesn't want to give me the five guilders"
(court testimony from Rubashov 1929: 158, date: 1465; also cited in Santorini 1992)

(2) zeyt gibetin d[a]z **mir** eyer fatr moykhl iz
be prayed that me-DAT your father forgiving is
"Hope/ask that your father is forgiving me"
(letter in Weinryb 1937: 54.6, date: 1588)

(3) d[a]z **es** unzr her gut oyz ginumn hut far an
that it our lord good out-took has presently
"...that our good Lord has made a success of it presently"
(Leib bar Moses Melir's *Book of Esther*, date: 1589)

In each of these clauses a weak object pronoun has scrambled to surface between the complementizer and the subject, in a position that is plausibly clause-adjoined under the standard assumption that the subject has moved to Spec(TP). However, it is possible to pinpoint the position of the scrambled element more definitively with the help of examples such as (1) above. In (1), the subject *yusf* appears to the left of sentential negation (and indeed, to the left of another object – a fact I return to below). Under the assumption that sentential negation appears to the left of vP[15], the subject must have moved out of its base position (to the right of negation) to some specifier above negation, Spec(TP) in the fairly restrictive phrase structure above. If this is correct, then the

[15] I do not take a stand on the exact position and status of negation in early Yiddish at this time.

scrambled pronoun must occupy an even higher structural position, which I take to be TP-adjoined. Another example of this type is (4) below:

(4) …dz **es** di mtsreym nit zaltn zehn
that **it** the Egyptians not should see.
"That the Egyptians shouldn't see it."
(Leib bar Moses Melir's *Book of Esther*, date: 1589)

Here again, the subject *di mtsreym* has moved to the left of sentential negation, *nit*, and the object pronoun *es* has moved farther up to clause-adjoined position.

Note that example (1) also clearly shows that early Yiddish had scrambling to multiple landing sites at different levels of the structure. As I mentioned above, the position of sentential negation *nit* signals the left edge of vP, and so both the indirect object *mir* and the direct object *di h' zhubim* ("the five guilders") must have scrambled out of their base object positions to the right of negation. The fact that the subject *yusf* can occur between the two scrambled objects plainly shows that there were at least two adjunction sites for scrambling in early Yiddish. Additionally, clauses like the one below argue that these two scrambling positions in early Yiddish were indeed TP and vP.

(5) Vi mir **das kinigreykh fun hkb"h un zeyn yokh** oyf uns antpfngn habn
how we **the kingdom of God and his yoke** on us accepted have
"…how we accepted upon ourselves the yoke of the kingdom of heaven"
(*Isaac ben Aaron Prossnitz's Preface to Sefer shir ha-shirim*, date: 1579)

In this sentence, both the subject and object appear to the left of the vP-modifying PP *oyf uns* (underlined above). The 1pl subject pronoun, *mir* (not to be confused with the

homophonous 1sg dative pronoun) is high in the structure in Spec(TP); indeed, Haeberli (2002) argues that subject pronouns in Germanic never remain in a low subject position, even when such a position is available for nominal subjects (as it seems to be in Old English). The object, *das kinigreykh*…, on the other hand, has scrambled from its base position past the vP-adjoined PP, but to a position below Spec(TP). This sort of example is further confirmation that there was a lower and a higher landing site for scrambled objects in early Yiddish.

3.3 Phrase Structure Variation and Change

As Yiddish changed over time from a German-style Tense-final clause structure to a Tense-medial clause structure, objects were trapped in a lower structural position due to the presence of the finite verb in Tense; they could no longer move to the high scrambling position adjacent to C. As I have already mentioned, the lower structural position for scrambling in modern Yiddish can be seen clearly be comparing the German sentences above to the parallel modern Yiddish sentences (repeated below as 6 and 7, respectively).

(6) a. Ich weiß daß **es** Johann gestern nicht gekauft hat.
I know that **it** Johann yesterday not bought has.

b. Ich weiß daß Johann **es** gestern nicht gekauft hat.
"I know that John didn't buy **it** yesterday."

(7) a. Ikh trakht az Hayim hot **im** nekhtn nit gekoyft.
I think that Hayim has **him** yesterday not bought.
"I think that Hayim didn't buy it yesterday."

b. Ikh trakht az Hayim hot nekhtn **im** nit gekoyft.
c. * Ikh trakht az Hayim **im** hot nekhtn nit gekoyft.
d. * Ikh trakht az **im** Hayim hot nekhtn nit gekoyft.

This is the expected effect of the change in the position of Tense on scrambling in Yiddish according to the GHC. However, the GHC actually makes a much stronger prediction: scrambling should be restricted to below Tense in **every** Tense-medial clause, at any time during or after the change in phrase structure, regardless of the overall frequency of Tense-medial vs. Tense-final clauses in the population at a given time. Using a parsed diachronic corpus of Yiddish (the *Penn Yiddish Corpus*, Santorini 1997/2008), this hypothesis can be tested in two ways. Note that all of the data below is taken solely from subordinate clauses, as in Santorini (1992, 1993a), since even Tense-final Yiddish matrix clauses were V2, exhibiting general verb movement to C as in German (following the standard analysis of West Germanic V2 going back to den Besten 1983).

First, the analysis of individual example sentences above demonstrated that early Yiddish had (at least) two potential landing sites for scrambled objects, adjoined to vP and adjoined to TP. During the time period when the change to Tense-medial was in progress, both Tense-final and Tense-medial TPs were produced by speakers and recorded in the texts that make up our data set. Without any *a priori* knowledge about how scrambling was restricted, one might make the maximally simple prediction that the

position of Tense in a given clause and the position of a scrambled object in a given clause are entirely independent of each other. If this were the case, then during the period of variation as the change in Tense was progressing, we would expect to see (subordinate) clauses of the following 4 types, i.e., all of the logical combinations of the possibilities for scrambling and for the position of Tense:

1. Tense-final, scrambling to vP-adjoined position
2. Tense-final, scrambling to TP-adjoined position
3. Tense-medial, scrambling to vP-adjoined position
4. Tense-medial, scrambling to TP-adjoined position

Statistically speaking, if any of the four combinations above were significantly favored or disfavored, it would call into question the independence of scrambling and the position of Tense. However, the GHC actually makes the more ambitious prediction that pattern #4 should be completely absent from the data set.

As it turns out, this is indeed the case: scrambling above Tense is entirely unattested in unambiguously Tense-medial clauses; in contrast to combination #2, combination #4 is impossible. As I mentioned above, due to the frequent use of extraposition in historical Yiddish and German, the majority of clauses in the corpus are ambiguous, in principal, and could have been generated by an underlying Tense-medial or Tense-final structure (e.g., the configuration **Subject > finite-lexical-verb > Object** could represent a Tense-medial clause or a Tense-final clause with object extraposition).

In order to address this issue, Santorini (1989; as well the other studies cited above) determined that certain light elements do not extrapose in Germanic, and so the following set of diagnostics signal that the finite verb has unambiguously moved to a medial Tense head: negation following the finite verb, verbal particles (separable prefixes) following the finite verb, and object pronouns following the finite verb. On the other hand, pre-finite-verb negation and particles are diagnostic of underlying Tense-final structure, as in examples (1)-(4) above.[16] In the table below, I show the frequency of objects preceding finite verb in clauses with one of the above diagnostics for the position of Tense.

Table 1. Clauses containing an object and a subject[17] preceding the finite (lexical or auxiliary verb)

Diagnostic	Obj, diagn. > Vfin	Obj > Vfin > diagn.	Vfin > Obj, diagn.	Totals
Verbal particle	16	**0**	57	73
Negation	32	**0**	191	223
Pronoun object[18]	N.A. (16)	**0**	87	103

[16] I cannot, of course, assume *a priori* that preverbal object pronouns are diagnostic of Tense-final structure, since the hypothesis I am out to test is whether or not any objects can move across medial Tense. However, weak object pronouns following the finite verb are diagnostic of Tense-medial structure, as these elements are too prosodically light to extrapose in Germanic.

[17] This condition is meant to exclude examples in which an object has been topicalized to the left of the finite verb in an embedded clause, rather than scrambled. It is a well-known fact of modern Yiddish that Tense-medial Yiddish embedded clauses allow topicalization with V2 (cf. den Besten & Moed-van Walraven 1986, Diesing 1990), and topicalized objects are not relevant to the discussion of scrambling.

[18] The "16" is in parentheses because, as I mentioned above, I can't assume that a *preceding* pronoun is diagnostic of Tense-final structure since this precisely the hypothesis I'm testing. Also, I have excluded one obvious scribal error (for which the translation is obscure, in addition) from consideration:

i. vi mir hrab"d zi"l hut mir lkhlumin eyn giroymt
how me-DAT the Rabbi D. of-blessed-memory has me-DAT nothing?? granted
"…how Rabbi D., of blessed memory, has granted me nothing(?)"
(West Yiddish court case, date: 1665)

The sentence above is certainly not a relevant example, since the two instances of *mir* are not actually two objects, but rather one repeated for reasons that are unclear. The resulting sentence is ungrammatical in

Totals	48 (without the 16)	**0**	335	383 (without the 16)

The column of zeroes[19] in the center clearly shows the effect of medial Tense on TP-adjoined scrambling: it is categorically disallowed.

On the other hand, when the relevant diagnostics precede the finite verb, showing the clause to be Tense-final, there are 46 examples of objects also preceding the finite verb. Below are examples of the diagnostic patterns which we find in early Yiddish; the first two clauses below contain the verb-particle Tense-final diagnostic, and the last two

---

Yiddish and German, and the author/copyist clearly meant to write only a single *mir*. Of course, whether the *mir* was originally intended to occur before or after the auxiliary is impossible to determine.

[19] This number excludes 2 examples that have negation following a finite auxiliary and preceding a nonfinite lexical verb which are most likely examples of negation trapped by verb-projection-raising (of the West Flemish type), from a Tense-final structure, rather than actual Tense-medial clauses. In these examples, the scope of negation is plausibly not sentential, but rather over the nonfinite verb; this is consistent with the fact the verb-projection-raising creates a scope island in the raised projection (Haeberli & Haegeman 1998). The examples are listed below:

ii. ven er zikh shun hut nit gimiat dribr…
If he REFL already has not had-trouble there-over
"If he hadn't already had trouble with it…"
(Preface to Jacob ben Isaac Rabbino Ashkenazi's *Preface to Sefer ha-Magid*, date: 1600)

In the above example, the scope of negation is ambiguous, as it generally will be in a past perfect sentence. However, the placement of the adverb *shun* before the auxiliary makes it very unlikely that this sentence is Tense-medial (this sentence is ungrammatical in modern, Tense-medial Yiddish), and so I assume that the negation scopes low.

iii. …dz mn mikh fr eyn krbn vil nit bgerin
that one me for a martyr wants not desire
"[what kind of sinner must I be] that people want to not desire me as a martyr"
(*The Vilna blood-libel case of 1690*, date: 1692)

In this case, the context makes it likely that the scope of negation is over *bgerin* ("desire"): this sentence occurs in a story in which a community makes a choice to redeem a prisoner against his will, and so it seems that the emphasis is on the community willingly ignoring his desire to be a martyr. But independent of the scope facts, the presence of the PP *fr eyn krbn* ("for a martyr") preceding the auxiliary makes it almost certain that this clause is Tense-final; this sentence would not be possible in modern Yiddish, nor would it be possible in other modern Tense-medial languages with V-to-T movement (e.g. Icelandic).

examples contain the negation Tense-final diagnostic. (The object is in boldface and the diagnostic is underlined.)

(8) das ikh **im** <u>ab</u> zag
that I him off spoke
"that I refused him"
(G"otz fun Fiderholtz's *Complaint*, in Birnbaum 1979: 159-160, date: 1518)

(9) ven ir **mir** **meyn zun** <u>um</u> brengt
if you me-DAT my son on bring
"if you kill my son"
(*Magen Avraham*, date: 1624)

(10) meyn kleydr hut er mir vr koyft di er **mir** <u>nit</u> hut gimkht
my clothes has he me-DAT sold which he me-DAT not has made
"and he sold me clothes that he hadn't made"
(court testimony from Rubashov 1929: 158, date: 1465)

Note that in the last example, the order of auxiliary and nonfinite verb in the relative clause is the same that a Tense-medial clause would have, but in this case, the diagnostic makes it clear that this is an instance of a Tense-final clause with the West Germanic Verb Raising construction (as in Dutch, West Flemish, or Swiss/South German). It is for this reason that I could not include the relative order of verb and auxiliary as a Tense-medial diagnostic in the table above. The example below, on the other hand, also shows the negation diagnostic, and additionally it has the expected order of verb and auxiliary for a Tense-final clause. (Note that this example also shows PP-extraposition, a commonly excercised option even in early Yiddish Tense-final clauses, as I mentioned above.)

(11) dz mn **di yungin** in ir yugnt nit givint hut tsu keynn peulut
that one the youngsters in their youth not won has to no occupation
"…that one hasn't gotten an occupation for his/her children in their youth"
(*Anshel ben Joseph's Preface to the Hebrew-Yiddish Bible Concordance*, date: 1534)

In sum, the diagnostic elements are evidently not evenly distributed among clauses containing objects to the left of the tensed verb, and so there must be an interaction between the choice of a given clause structure (Tense-medial vs. Tense-final), which the diagnostics represent, and the position of a scrambled object. If an object originates to the right of Tense, it cannot scramble to a hierarchical position above Tense. Thus, the gap in the historical data corresponds to the negative judgments shown above in (7) for speakers of modern Yiddish.

Aside from clauses with the diagnostic elements listed above, the only other unambiguous clause type in the corpus (i.e. unambiguous without making reference to object position) are clauses with auxiliaries and the order, "nonfinite-lexical verb > finite-auxiliary," which are unambiguously Tense-final regardless of object placement (see 1, 3, 5, 11 above). In contrast, most clauses during the period of phrase structure change in Yiddish are stringwise ambiguous between Tense-medial and Tense-final from the point of view of a modern analyst reading the texts. The remaining types of clauses either have an auxiliary and the order, "Aux > V", which could in principal be either Tense-medial, or Tense-final plus West Gmc. Verb-Projection-Raising, as in the sentences below:

(12) zi **zal** eydes **zagn**
She should testimony say
"She should give testimony"

(West Yiddish court case, date: 1565)

(13) dz shlmh hmlkh eliu hshlum **hat** shir hshirim **gimkht**
that Salomon the-king peace-be-upon-him has *Song of Songs* made
"…that Salomon The King, peace be upon him, wrote *The Song of Songs*"
(*Isaac ben Aaron Prossnitz's Preface to Sefer shir ha-shirim*, date: 1579)

And finally, there are clauses with a finite lexical verb and no auxiliary; these contain no hint to their underlying structure, minus any information about how object scrambling is constrained (i.e. without the GHC, the first clause below must be considered as ambiguous as the second clause below).

(14) ven mn eynm kind afilu gibt eyn shtuk brot
if one a child even gives a bit bread
"if one even gives a child a piece of bread."
(*Isaac ben Aaron Prossnitz's Sefer shir ha-shirim*, date: 1579)

(15) dz zi lubn hkbh
that they love God
(*Isaac ben Eliakum's Preface to Lev Tov*, date: 1620)

The potentially ambiguous clause types cannot be directly used to answer the question of whether the change in the position of Tense had an effect on scrambling. However, we can make indirect use of this data by setting up the following quantitative experiment.

Beginning with Kroch (1989), Santorini (1989), Pintzuk (1991), and continuing in much subsequent work, it has been found that during periods of change in a given syntactic parameter, different constructions in a language that represent instances of that parameter underlyingly (different contexts for the parameter) will increase in frequency or decline in frequency at the same rate; this is the "Constant Rate Effect" (Kroch 1989).

Taking the Constant Rate Effect as an empirical fact, different types of Tense-final clauses should decline in frequency over time at the same rate, as the various types of Tense-medial clauses rise in frequency at the same rate. By hypothesis, all clauses with objects preceding the finite verb are Tense-final; if the GHC is valid, then the appearance of an object to the left of a finite verb should itself diagnose Tense-final structure. This gives us the following set of clause types containing at least one object, and the underlying structures they are either known or hypothesized to represent.

Table 2.

Clauses with Auxiliaries		
Tense-Final		**Tense-Medial**
Obj > V > Aux (known Tense-final)		Aux > Obj > V (hypothesized/plausibly[20] Tense-medial)
V > Aux > Obj (known Tense-final)		Aux > V > Obj (hypothesized/plausibly Tense-medial)
Obj > Aux > V (hypothesized Tense-final)		
Clauses with a Finite Lexical Verb (no auxiliary)		
Obj > finite-lexical-Verb (hypothesized Tense-final)		finite-lexical-Verb > Obj (hypothesized/plausibly Tense-medial)

According to the Constant Rate Effect, since there was a general change in Yiddish from Tense-final to Tense-medial, the hypothesis of the GHC will be considerably

[20] These clauses are plausibly Tense-final, but they could theoretically also be Tense-final with West Germanic verb-raising plus object extraposition. However, the different types of plausibly Tense-medial clauses that are compared in the experiment below all contain some Tense-final clauses with extraposition and/or verb-raising, so the comparison between types is valid (Tense-medial is slightly undercounted across the board). This is the reason that these clause types are termed "hypothesized/plausibly, rather than simply "hypothesized"; the "hypothesized" Tense-final types are predicted by the GHC to be pure collections of Tense-final clauses. As it turns out, the interference from extraposition and verb-raising in the plausibly Tense-medial types actually is mild enough that the result is still clear despite this obvious source of noise in the data.

strengthened if the clause types that are hypothesized (or known) to be Tense-final decline at the same rate, and the clause types that are hypothesized to be Tense-medial increase in frequency at the same rate.

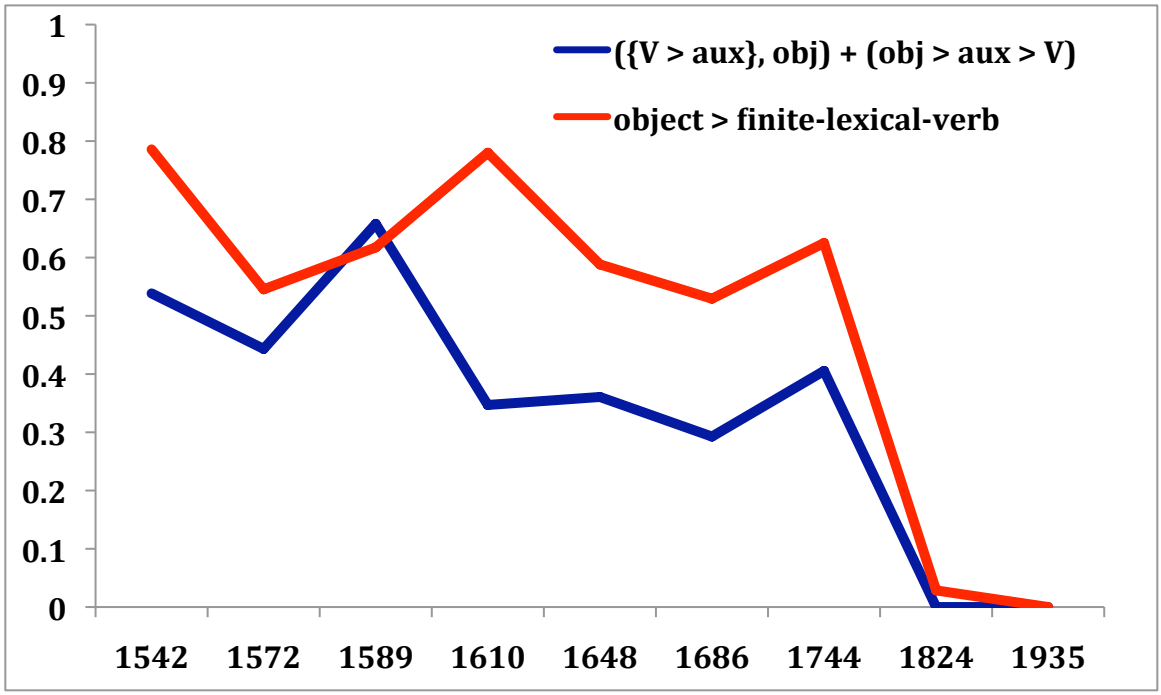
Ideally, we would test the connection between the different sentence types with the Constant Rate Effect by checking if each known and hypothesized Tense-final type declines at the same rate as every other one (and correspondingly, that every Tense-medial type increases in frequency at the same rate as every other). Unfortunately, splitting up the data in this way, by this many clause types and by time periods (of any reasonable granularity), results in numbers that are so small that the natural noise in the system obscures any overall trend over the time variable. (And even without such a fine-grained division, the data are still noisy, as will be apparent below.) This being the case, the next best test is to group the data in a reasonable way and compare groups that are hypothesized to have the same behavior over time. For the purposes of this experiment, I divided the data into subordinate clauses containing both objects and auxiliaries and subordinate clauses containing objects and a finite lexical verb, as in Table 2 above. Within each clause type, I grouped the types with auxiliaries which are hypothesized to have a given phrase structure, and then compared them to the respective type with a finite lexical verb that is hypothesized to have the same phrase structure. If the GHC is a valid descriptive generalization, the hypothesized Tense-final types with auxiliaries should decline at the same rate over the history of the language as the hypothesized Tense-final type with finite lexical verbs, in comparison with the respective hypothesized/plausible Tense-medial clause types.

In a sense, this is a very rough type of quantitative experiment. If the relevant groups of clause types do not decline in frequency at the same rate, then in fact, very little can be concluded. There are many ways in which these clause types can differ from each other, in addition to the one way under investigation, and it is not possible to control for all of them with the available data set. From this perspective, it would be unsurprising if the decline of hypothesized Tense-final clauses with auxiliaries and the decline of hypothesized Tense-final clauses with finite lexical verbs did not show any striking similarity. This would not necessarily contradict the GHC, as the effect could easily be due to some other (not necessarily syntactic) factor, but in such a case we would learn nothing. On the other hand, if the comparison is interpretable and matches the prediction of the GHC, in spite of the inherently noisy nature of this type of historical data, then it is surprising, and surprising in an instructive way. In the event that the theory predicts a real quantitative effect, one which is strong enough to emerge out of the considerable noise in the system, the complications in the experiment actually come to tell in favor of the hypothesis instead of detracting from it. This type of somewhat indirect experimentation is necessary in face of the type of documentary data set that historical linguists are frequently faced with. It is a simple attempt at what Labov (1994: 11, and p.c.) calls, "making the best use of bad data." There are many reasons why the experiment might not work, but if it does, it can be taken together with the other data in

support of the GHC, and then each piece of data complements and strengthens the impact of the others.[21]

Fortunately, in this case there was a clear quantitative effect, and it supports the GHC. The auxiliary group and the finite lexical verb group which are hypothesized to be underlyingly Tense-final decline at the same rate, and the corresponding Tense-medial types rise over time at the same rate. The results of the study are shown in the charts and tables below.[22]

Figure 2. Decline in Hypothesized/Known Tense-final Clause Types



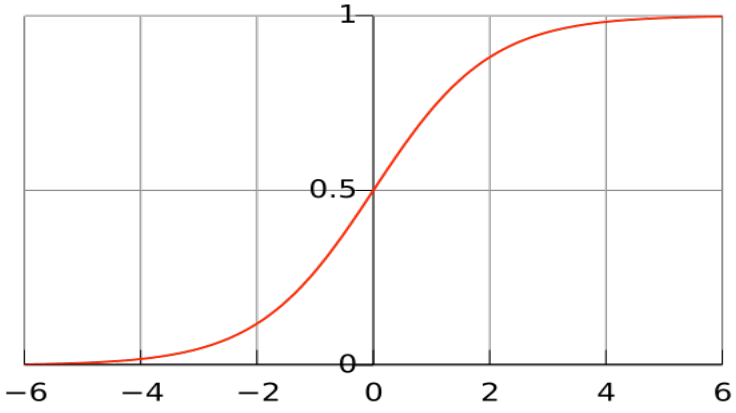
[21] Note that in the *Penn Yiddish Corpus* (Santorini 1998/2008), there is no annotation to distinguish pronominal objects from nominal DP objects, so I have included both types in all counts. While this is not ideal, it has clearly not obscured the effect shown below.

[22] In order to control for the effects of dialect over time, this study used data exclusively from East Yiddish texts. Another reason for this choice is that West Yiddish texts do not have the same time-depth as East Yiddish texts, since West Yiddish became extinct (at least, in written records) sometime in the 18th century.

By merely eyeballing the curves shown above, one can see that they are close to being the same curve, but we can arrive at a much more impressive result by fitting the logistic function to the data in each case and comparing the two models. The tried-and-true method for determining and comparing the rate of change of different syntactic variants is to model each change with the logistic function (shown below with its plot to the right), and then to compare the estimated slope parameters for the contexts that are hypothesized to be reflexes of a single syntactic change (see Kroch 1989, Santorini 1993a, Pintzuk & Taylor 2004, Pintzuk 2005, and references therein for examples of this use of the logistic). Of course, modeling dynamic systems with the logistic function is by no means specific to language change: a logistic curve is the characteristic shape of evolutionary competition between two variants in a population with finite resources (Nowak 2005: 12).

Figure 3. Basic Logistic Function

$$P(t) = \frac{e^t}{1 + e^t}$$

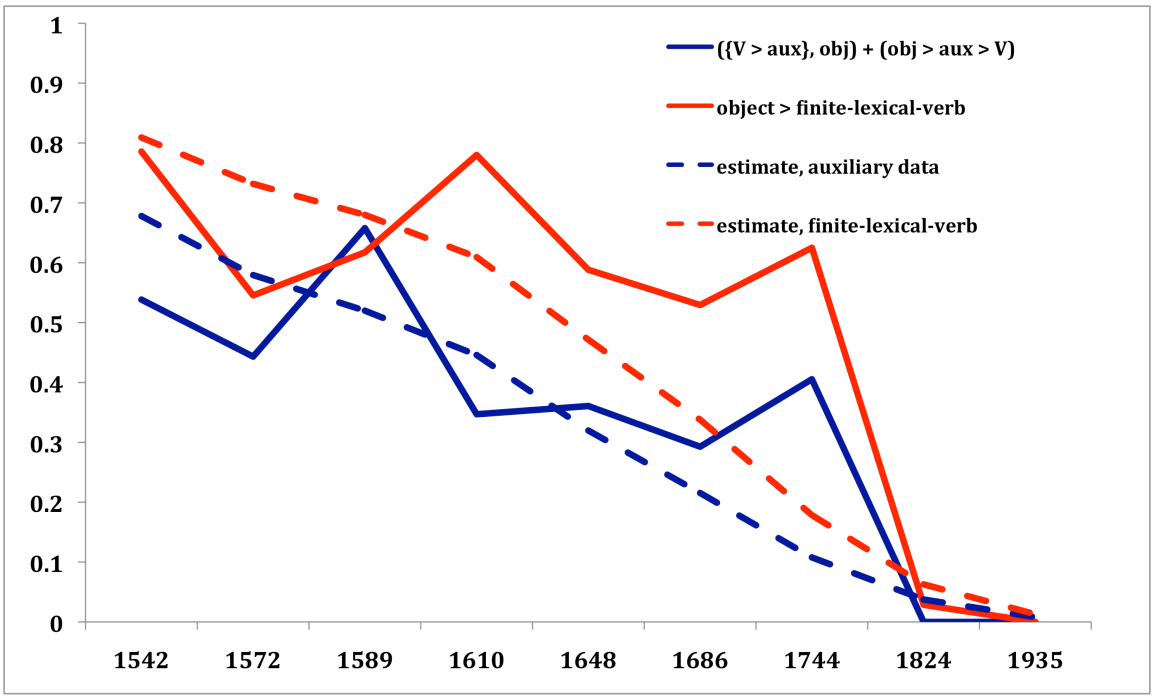


This is the simplest form of the logistic, plotting probability (P) over time (t) (the corresponding linear transform of the logistic, the logit, in this case has a slope of 1 and a y-intercept of 0). However, it is possible to define different logistic curves with different slopes and positions along the x-axis by replacing *t* in the equation above with $k+rt$. *k* is a constant determining the position of the curve along the x-axis (the y-intercept of the corresponding logit), and in linguistic terms, the x-axis position of the curve models the point in time the linguistic change began and the time-range of the change. *r* determines the slope of the curve (*r* is the slope of the logit), which in linguistic terms refers to the rate of change at which the new variant replaces the old one in the speech community.

Any change in which one linguistic variant replaces another in a finite population can be given an idealized model (i.e. controlling for random noise) by fitting the logistic function to the observed data, arriving at the best fit by allowing the *k* and *r* parameters to vary. After estimating these parameters for each set of data, it becomes possible to model the decline of each set of clause types with the same equation and then to compare their diachronic developments directly. When two different diachronic developments have best-fit logistic models with the same (i.e. not significantly different) slope parameters, the Constant Rate Effect has been found, and it is grounds for supposing that the two different diachronic developments are in fact reflexes of a single underlying change.

The S-shaped curves produced by the logistic estimate of the change for the auxiliary and finite lexical verb cases are shown in Figure 4 below (along with a plot of the actual frequencies, repeated from above).

Figure 4. Decline in Hypothesized/Known Tense-final Clause Types, with Logistic Models



The estimated logistic/logit model parameters are shown beneath the figure above, and the frequencies and estimates from the logistic models for each type are shown in the tables below.

Table 3. Subordinate Clauses with Objects and Auxiliaries (frequencies and estimates shown)

DATE	obj > finite-V	finite-V > obj	Totals	frequency: obj > finite-V	logistic estimate: obj > finite-V
1534-1550	**11**	**3**	14	0.786	0.809
1565-1579	**24**	**20**	44	0.545	0.732
1588-1590	**92**	**57**	149	0.617	0.680
1600-1620	**39**	**11**	50	0.780	0.609
1624-1671	**10**	**7**	17	0.588	0.471
1675-1697	**27**	**24**	51	0.529	0.338
1704-1783	**5**	**3**	8	0.625	0.179
1800-1848	**1**	**34**	35	0.029	0.063
1910-1959	**0**	**373**	373	0.000	0.013

Table 4. Subordinate Clauses with Objects and Finite Lexical Verbs

DATE	{(mv > aux),obj} **+** (obj > aux > mv)	(aux > obj > mv) + (aux > mv > obj)	Totals	frequency: Tense-final (hyp.)	logistic estimate: Tense-final (hyp.)
1534-1550	**7**	**6**	13	**0.539**	**0.678**
1565-1579	**39**	**49**	88	**0.443**	**0.579**
1588-1590	**50**	**26**	76	**0.658**	**0.519**
1600-1620	**34**	**64**	98	**0.347**	**0.446**
1624-1671	**22**	**39**	61	**0.361**	**0.319**
1675-1697	**36**	**87**	123	**0.293**	**0.215**
1704-1783	**15**	**22**	37	**0.405**	**0.108**
1800-1848	**0**	**151**	151	**0**	**0.0378**
1910-1959	**0**	**551**	551	**0**	**0.008**

It is clear from the slope (r) parameters shown under Figure 4 that the two hypothetically Tense-final groups decline at the same rate over time, indicating that they are reflexes of the same change: the loss of Tense-final phrase structure in Yiddish. According to the best models for the frequencies over time of each group of clause types, the slopes for the hypothesized Tense-final auxiliary case and the hypothesized Tense-final finite-lexical verb are identical to 3 decimal places, showing that the decline in frequency of the two groups occurs at the same rate over this large time period. However, this can be made even more precise with a simple test to determine if the two estimated slopes are significantly different from each other. The test proceeds as follows:

re-fit the logistic to both the auxiliary and finite lexical verb data, but this time constraining both models to use a single slope parameter (*r* is the same). Then compare the total log-likelihood (a measure of the model's fit to the data) for this new, combined model, with a sum of the log-likelihoods of the two original models, in order to determine whether the combined model fits the data in a significantly different way from the two original models. In other words, if we model the two changes with a single slope rather than with two curves with different slopes, does the model fit the data significantly worse, or is the fit about the same as before?

In this case, the comparison between the combined-slope model and the two original models (the log-likelihood ratio test statistic; cf. Agresti 1996:110) evaluated against a chi-square distribution with 1 degree of freedom gives a p-value of .73; the slopes of the two curves are not significantly different given the observed data, and in fact, if one assumes tha the curves have the same slope, there's a 73% chance that the observed difference between the data would happen by chance. Furthermore, even the intercept (k) parameters for the two models are quite close. This makes it at least possible that the two cases are even more similar than the Constant Rate Hypothesis (Kroch 1989) would predict: the two sets of syntactic patterns leave the language not only at the same rate, but quite possibly along exactly the same trajectory (modulo the noise in the system, which the logistic models are able to abstract away from).

The fact that this quantitative study showed such a robust pattern, despite the unavoidable noise, constitutes another type of evidence that object position (i.e. scrambling) is constrained by the presence of left-headed functional projections. The

position of Tense determines the scrambling possibilities, and so it is possible to observe the Tense-final to Tense-medial change in Yiddish simply by observing the position of objects relative to the verb (in non-topicalization contexts). This result also represents a methodological advance, because we can now add to the set of diagnostics for identifying Tense-final clauses in historical texts from languages with phrase structure variation: minus the effect of some non-scrambling movement process (e.g. A'-movement or head-movement), the presence of a preverbal object diagnoses a head-final structure.

Note also that the change apparent from the charts and tables above cannot be explained as an independent change in the nature of scrambling in Yiddish; that is, it is not possible to argue that the GHC is irrelevant here and clause-adjoined scrambling was independently lost during the history of Yiddish. Since the Tense-final to Tense-medial has been independently verified in previous work (Santorini 1992, 1993a), any additional change would have to add to (or detract from) the effect of the change in Tense on the frequencies of the different clause types. For example, compare the trajectory of the change for the finite lexical verb case above to that of the auxiliary case. If the change in object position over time in the finite lexical verb case (loss of the Obj > V-finite pattern) were primarily due to a change in scrambling, which happened to overlap with the change in Tense, then the shape of that curve should be different from the shape of the curve in the auxiliary case; the auxiliary case includes a number of unambiguously Tense-final clauses (the V > Aux cases) in calculating the overall frequency of hypothesized Tense-final clauses. Since the V > Aux clauses are unambiguously Tense-

final, they could not derived from scrambling across Tense even under a theory in which such scrambling is possible. Because of the inclusion of these Tense-final types, the auxiliary case should have an artificially inflated number of hypothesized Tense-final clauses at the beginning of the change, but not at the end, at which point Yiddish has become uniformly Tense-medial and also does not allow scrambling past Tense.

For this reason, under the theory that there were two overlapping changes, the curve for the auxiliary case should be steeper than that of the finite lexical verb case, showing too high a frequency for Tense-final at the beginning of the change and then dropping off more sharply than the lexical verb case at the end. Instead, what the data actually shows is the same slope for the two cases, and if anything, the slope for the auxiliary case is very slightly more gradual (-.0142) than the finite lexical verb case (-.0147), not the other way around. The data show that in comparing the hypothesized Tense-final and hypothesized Tense-medial clauses in the auxiliary and the finite lexical verb cases, we are not comparing apples and oranges, and the simplest explanation for the comparable behavior of these cases is that it's a single force which drove the change in frequency in both cases.

Thus, the diachronic path of Yiddish shows us that the modern Yiddish scrambling system is exactly the same as the modern German scrambling system, but with a single parameter toggled. The Generalized Holmberg Constraint does not change, but the change in the Tense parameter created a new context where it could assert itself. Note that this result is virtually identical to the findings of Taylor's (1990) study of the relationship between "clitic" position and clause-structure change in Pre-Medieval Greek.

(I use the non-traditional term "Pre-Medieval" in order to avoid confusion: the Taylor (1990) considers Ancient Greek from the early, Homeric period through Koine New Testament Greek. Her study did not include the Medieval Greek dialects described in Condoravdi & Kiparsky 2002, 2004)[23]. Like Yiddish, Pre-Medieval Greek changed from Tense-final to Tense-medial, and it also had a system of weak pronoun scrambling ("clitics" in the traditional Greek terminology). In unambiguously Tense-final clauses, Greek clitics show behavior that is quite similar to that of weak object pronouns (and other scrambled elements) in German or the other OV Germanic languages: they scramble leftward potentially as far as C, left-adjoining to maximal projections (as Taylor discusses, this is the simplest analysis of the syntax once the effects of prosodic inversion

[23] Condoravdi & Kiparsky (2002, 2004) provide a comprehensive analysis of clitic pronoun position in the Late Medieval Greek dialects and a representative sample of the modern Greek dialects. Their results suggest that clitic behavior in most, if not all, of these dialects is not object scrambling of the Germanic kind, but rather true clitic behavior with each dialect showing slightly different prosodic and syntactic conditions on clitic placement. However, it is not necessarily the case that their analysis is incompatible with Taylor's results, even though the latter receive their simplest explanation if Ancient Greek clitics were scrambled objects which responded over time to the change in the position of Tense. It is possible that the situation in the Late Medieval and modern dialects described by Condoravdi & Kiparsky (2002, 2004) was the result of later developments in the clitic syntax of those varieties, and/or contemporary developments in Koiné Greek that were not represented in the New Testament.

In fact, Condoravdi & Kiparsky (2002: section 5.3) claim that the Pontic dialect is descended from a variety in which "clitics were predominantly postverbal," which may in fact be the outcome of the changes that Taylor (1990) describes, minus any later changes in the status of the classical Greek clitics/weak pronouns. Condoravdi & Kiparsky suggest that a postverbal pronoun system may have already begun to arise in the classical Greek period, citing as evidence the fact that Herodotus shows variation in pronoun placement, includes postverbal pronouns; this is exactly the evidence that Taylor identifies as the starting point for the Tense-final to Tense-medial change in pre-Medieval Greek, and the consequent change in pronoun position (see the examples I cite from Taylor below). Condoravdi & Kiparsky (2002: 25) also note that the medieval Pontic of the monestary of Vazelon shows postverbal pronouns in contexts where the Late Medieval Greek dialects cannot have postverbal pronouns (though they all have postverbal pronouns in certain contexts), which indicates that medieval Pontic had a more uniformly postverbal clitic/pronoun syntax. If this is correct, then medieval Pontic may have represented a later stage of the changes that Taylor (1990): Tense-medial, but with simple weak pronoun scrambling rather than a true clitic system. Also, if Condoravdi & Kiparsky (2002, 2004), taken together with Taylor (1990), have indeed identified cases in Greek in which true clitic systems emerge from weak pronoun scrambling after/during a period of phrase structure change, then the developments in the Greek dialects might be very similar to the development of a true clitic system in Old and Early Middle English which I argued for in Wallenberg (2008), and which I argue for in Chapter 9 of this dissertation.

are accounted for). In fact, Taylor shows that the final landing site of scrambled clitics can be any one of a number of phrasally-adjoined positions, in the case that TP and/or vP are recursive due to the presence of adjoined modifiers or the scrambling of other elements in the clause. Below I repeat two of Taylor's examples, showing two landing sites for clitic pronouns in predominantly Tense-final Homeric Greek.

(16) aiei =**toi** muthoi philoi akritoi eisin
always to-you words dear endless are
"endless words are always dear to you"
(Illiad, 2.796, Taylor 1990: 126)

(17) akhlun d' au =**toi** ap' ophthalmo:n helon
mist P again your from eyes took-away
"I took the mist away from your eyes again"
(Illiad, 5.127, Taylor 1990: 127)

The change in the position of Tense in Greek can be seen in the loss and eventual disappearance of Tense-final clauses such as (18) below, and the rise of clauses with the Tense-medial pattern in (19). By the time of New Testament Greek, the Tense-medial system has become far and away the dominant system. The two clauses below also each contain a clitic pronoun, which I have put in boldface.[24]

(18) kai =**min** deutero:i etei talantou Aigine:tai de:mosie:i misthountai
and him in-second year talent Aeginetans for-public-service paid
"and in the second year the Aeginetans paid him a talent for public service"
(Herodotus 3.131.2, Taylor 1990: 145)

[24] Following Taylor (1990), from which these examples are taken, I use a Romanized orthography rather than traditional Greek orthography, with the only difference being that I indicate long vowels with following colons rather than bars. As in Taylor, "=" indicates a phonological clitic-host boundary.

(19) ho arkhiereus ekse:neike =**sphi** phialas khruseas
the high-priest brough-out to-them cups gold
"the high priest brought gold cups out to them"
(Herodotus 2.151.1, Taylor 1990: 147)

Apart from the clitic/weak pronoun in (19), this latter type of sentence (surface SVO sentences) is ambiguous with respect to the position of Tense, because of the possibility that it was derived by rightward extraposition of the object across the verb. Even so, as Taylor discusses, some of the ambiguous sentences are potentially Tense-medial even in the earlier stages of the language in which clauses are predominantly Tense-final. The frequency of these ambiguous clauses increases (along with that of other head-initial structures in the language) as the change progesses from Homeric Greek through New Testament Koine Greek, in roughly 100 C.E. Unambiguously Tense-final clauses (with multiple heavy constituents preceding the finite verb; cf. Pintzuk 1990, Santorini 1992, Santorini 1993 for further discussion of this diagnostic), such as (18), steadily decrease in frequency and ultimately disappear from the language.

But it is the position of clitic arguments in sentences like (18) and (19) that is the crucial point for the purposes of the current study. Taylor shows that as the frequency of the ambiguous, SVO clauses increases at the expense of unambiguously Tense-final clauses, the position of clitic pronouns also changes in the (otherwise) ambiguous SVO clauses. As the change to left-headed TPs progresses in Greek, weak pronouns cease to appear in preverbal positions, and increasingly appear in a position immediately following the finite verb. Furthermore, the post-tensed-verb clitic position is not available in unambiguously Tense-final Pre-Medieval Greek clauses (Taylor 1990:154),

indicating that the post-verbal position is purely a feature of the Tense-medial grammar. Again, the Generalized Holmberg Constraint predicts the change in clitic position that Taylor observed if Pre-Medieval Greek clitics were indeed weak pronouns with scrambling behavior, as Taylor's study suggests. Viewed from this perspective, both the change in Greek clitic position and the change in object scrambling in Yiddish are entirely reducible to the change in the position of Tense. If the Generalized Holmberg Constraint is simply a universal and, by definition, cannot change, then the development of a left-headed TP immediately inserts a barrier to leftward scrambling, and any elements that are allowed to scramble leftward in Greek and Yiddish are immediately trapped in lower surface structure positions than they were under the old Tense-final grammar.

3.4 What doesn't change, doesn't change: a remnant of TP-adjoined scrambling in modern Yiddish

In addition to the quantitative argument above, there is another reason to reject the idea that the diachronic devlopement of Yiddish involves an independent loss of scrambling to TP-adjoined position: in the relevant context, scrambling to TP-adjoined position was never actually lost at all. It is true that the change to Tense-medial restricted scrambling to positions below Tense in subordinate clauses (the data set I considered in the quantitative study), but this is something of a simplification where matrix clauses are

concerned. In root clauses, where V-to-C movement is a possibility, scrambling is restricted to below Tense just in case the tensed verb lands in Tense at the end of the derivation. However, in contexts where the tensed verb is attracted to C, an object can scramble as high as it could in any stage of early Yiddish: up to TP-adjoined position and potentially crossing the subject. The fact that this behavior persists in modern Yiddish shows that nothing has changed in the history of Yiddish concerning the scrambling operation itself. Rather the GHC is simply apparent in a larger set of contexts in modern Yiddish than it was under the old phrase structure.

The high scrambling position is shown in the following matrix clauses, all written in the 20th century (more than a hundred years after any potentially Tense-final examples can be found in even highly literary texts). Note also that *Royte Pomerantsen*, the text in which examples (20), (21), and (23) below occur, was written as a compilation of stories reflecting colloquial Yiddish as the author heard it (see the introduction in Olsvanger 1947), and so it is unlikely that these sentences represent an archaism of some kind or a feature of a highly formal register.

Scrambling to TP, under V-to-C movement

(The scrambled object in the sentences below is in boldface, and diagnostic elements are underlined.)

(20) hot **dos** di rebetsn nit gekent hern,
has **that** the rebbetzin not could hear
"The rabbi's wife wasn't able to hear that.
[she couldn't stand to hear it]"

(Olsvanger 1947, *Royte Pomerantsen*, token 243, date: 1947)

(21) Hot **im** der rebe gegebn an eytse, "du bist dokh an amorets"
Has **him-DAT** the rabbi given a piece-of-advice, …
"The rabbi gave him some advice, …"
(Olsvanger 1947, *Royte Pomerantsen*, token 235, date: 1947)

(22) farvos zoln **zikh** yidn glat krign?
why should **REFL** Jews in-general fight
"Why should Jews always fight amongst themselves?!"
(Perets Hirshbeyn, *Grine Felder*, token 103, date: 1910)

(23) nu, (iz) vos hot **aykh** der fish geentfert ?
So, what had **you-DAT-PL** the fish answered
"So what did the fish say back to you?"
(Olsvanger 1947, *Royte Pomerantsen*, token 47, date: 1947)

Note that each of these sentences contains not only a scrambled object to the left of the subject, but also some context that is a well-established V-to-C trigger across Germanic. The first two sentences above are narrative V1 sentences, generally taken to result from V-to-C movement triggered by an appropriate information-structural context; this is presumably encoded in the syntax as a feature on C and an empty operator in Spec(CP) (see Besten & Moed-van Walraven 1986 on Yiddish, Thráinsson 1986 on Icelandic, and Kroch & Taylor 1995 as well as references cited there on the set of V-to-C contexts in Old English). The last two sentences are verb movement triggered by wh-movement, an unconstroversial V-to-C trigger cross-linguistically (for languages with wh-movement, of course). Thus, the object scrambling in these cases is clearly licensed by V-to-T-to-C movement, regardless of whether one accepts the claim in Diesing (1990) that finite verbs in modern Yiddish typically remain low in Tense even in many root

contexts. Furthermore, examples (20) and (22) provide additional evidence from the position of the subject that the object has scrambled past the original position of Tense, following the incorporation of the verb with Tense and subsequent further movement up to C. In example (20), the subject appears to the left of sentential negation (underlined), which could not possibly be occupying a position any lower than the left edge of the vP projected by the auxiliary *gekent* ("could"); note that this is obvious independently of how one analyzes the order of the nonfinite verbs here, since the negation clearly scopes over the modal in context. The position of negation must be higher than the base position of the subject, and it may in fact be in a higher position than the left edge of vP since it potentially even has sentential scope. This indicates that the subject, *di rebetsn* ("the rabbi's wife"), has moved to Spec(TP), and the object, *dos* (demonstrative "that") has scrambled above Spec(TP), just as in early Yiddish subordinate clauses. Similarly, the subject *yidn* ("Jews") in (22) appears to the left of the adverb *glat* ("in general", "always"), which clearly has scope over the whole event including the agent. Thus, the subject has moved out of vP and plausibly to Spec(TP), and the object *zikh* (the reflexive, "themselves") must have scrambled out of TP. The position of the subject in these cases also makes it certain that the verb has raised higher than Tense, and given standard assumptions about verb raising in Germanic (including modern English), it is V-to-C movement that has licensed the object-scrambling in these cases.

I must also briefly note here that while this discussion has remained agnostic with regard to much of Diesing's (1990) analysis of the position of the subject and finite verb in Yiddish, the existence of examples such as (20) and (22) above is a serious problem

for two aspects of Diesing's analysis: first, the idea that "finite-V > Subject" always results from V-to-I movement over the subject, and secondly, that wh-words front to Spec(IP) in Yiddish, leaving the subject in Spec(VP) in matrix wh-questions ("IP" and "VP" corresponding to my "TP" and "vP," according to the terminology used in that paper). In both sentences, the subject has moved leftward across a modifier with potentially sentential scope or negation, neither of which can be lower than vP-adjoined in order to have scope over the entire event (and the modifier and the negation may, in fact, be adjoined even higher than vP). However, the subject cannot have moved to Spec(TP) in (22) if the wh-word appears in Spec(TP). On the other hand, this word order is entirely expected if the wh-word is in Spec(CP) and the subject is in Spec(TP), as is standardly assumed for other Germanic languages (and beyond). The only way Diesing could maintain that the subject is in its base position in Spec(VP) (or Spec(vP), depending on one's phrase structure) is to assume that negation and adverbs in Yiddish can occur below the subject's base position adjoining to (or taking as their complement) a non-maximal projection (v' or V'). Not only would this be undesirable from the perspective that it would allow modifiers to adjoin to non-maximal projections (or worse, allow heads to take non-maximal complements), it would not allow negation or adverbs in examples like (20) or (22) to scope over the subject's base position.

The basic analysis in Diesing (1990) could be salvaged by assuming that there is another available subject position in the specifier of some other "Split-Infl" (Pollock 1989) projection between TP and vP, but it is unclear if this modification would be constrainable in such a way as to be consistent with the data that led to Diesing's original

proposal. Under this modification of Diesing (1990), subjects could appear in Spec(TP) or in the specifier of Tense's complement (whatever that projection is), but they move out of vP/VP in either case. This lower subject position would need to be constrained to host only subjects, whereas Spec(TP) could host either subjects or non-subject topics, as in Diesing's original analysis. If the lower subject position is obligatorily a case checking position for nominative only (perhaps AgrSP, if a phrase structure with a low AgrSP is tenable) by the head-spec configuration, but Spec(TP) is not always a position for checking nominative case, then the former position could be restricted to subjects and the latter could host either subjects or non-subject topics (presumably with subjects moving through Spec(AgrSP) on their way to Spec(TP) in the case that they are topics), as in Diesing (1990). As long as sentential adverbs and negation were then allowed to appear below the low subject position, the substance of Diesing's (1990) analysis could still be maintained, including the movement of wh-words to Spec(TP). However, this analysis would still have the effect of making Yiddish odd cross-linguistically in not having operators move to Spec(CP) (and consequently triggering V-to-C movement), and it would also need to posit a non-trivial reanalysis of Yiddish clause structure in order to explain why V-to-C movement left the language. Like German, there was a clear main-subordinate clause asymmetry with respect to V2 in early Yiddish and no indication that V-to-C movement was lost in the language. It would also fail to connect Yiddish subordinate clause topicalization (and Icelandic embedded topicalization, if the same analysis were applied to both languages) to the types of embedded topicalization which are allowed in every other Germanic language, which

have been shown by Iatridou & Kroch (1992) to involve CP-recursion. However, I will leave this issue as a problem for further research.

In this way, the relationship between verb-movement and scrambling in Yiddish is exactly analogous to that of "object shift" in the Germanic VO languages (especially Swedish and Early Modern English, in which scrambling can target a TP-adjoined position under V-to-T-to-C movement – cf. Josefsson 1992 for Swedish and Wallenberg 2007 for EME). This is an extension of Holmberg's (1986) Generalization which is entirely predicted under the GHC. Even though there are a number of potential barriers to scrambling in the form of functional heads to the left of the scrambled item, head-to-head movement results in incorporation of the moved head into the target head (Roberts 2001), and it is the resulting complex head which constitutes a barrier to scrambling of the object. Thus, movement of Tense to Comp creates a configuration in which scrambling can proceed to the same position it normally targeted in early Yiddish.

Another parallelism between scrambling and Scandinavian object shift is illustrated by the example in (22). This sentence is especially significant in that the scrambled object is the reflexive, *zikh*, bound by the subject *yidn*. This type of example demonstrates that scrambling in Yiddish can reconstruct if the object is scrambled past the subject and the subject is a potential binder for the object; otherwise this sentence would constitute an obvious violation of Principle A. This property of scrambling past the subject has been well investigated in other scrambling languages, notably in German and Korean (Lee & Santorini 1994). Thus, the Yiddish example above is exactly parallel to the German example in (24) below, where the cognate reflexive *sich* has scrambled

past the subject up to a position adjacent to C, and it is able to reconstruct and be bound by the subject, *Johann*.

(24) Gestern hat **sich** Johann rasiert.
Yesterday has **REFL** John shaved.
"Yesterday, John shaved himself."

A less well-known fact is that in Swedish, typically considered an "object shift" language rather than a scrambling language, one can see that object shift shows the same reconstruction effects as in German or Korean.[25] This is shown by the position of the reflexive pronoun *sig*, which occurs to the left of the subject in the Swedish examples below.

(25) Klarar **sig** barnen på egen hand?
manage **REFL** children-the on own hand
"Do the children manage on their own?"
(Holmberg 1984: 3)

(26) Igår lade **sig** mamma tidigt.

[25] Sentences in which objects are scrambled past the subject are generally judged ungrammatical by speakers of Scandinavian languages other than Swedish, so this question cannot be tested in these varieties. It is unclear why Swedish should be alone in Scandinavian in allowing this type of scrambling/object shift, though it is important to note that judgments are frequently unclear and can differ from speaker to speaker, particularly for Icelandic (see Hellan & Platzack 1995: 51); for example, Thórhallur Eythórsson (p.c.) reports the following example to be grammatical, though he does not generally accept this type of sentence:

iv. Í gær kysstu hana allir.
Yesterday kissed her-ACC all-NOM
"Yesterday everybody kissed her."

An Icelandic sentence like this might represent scrambling past the subject under V-to-C movement, as in the Swedish case, but it is actually ambiguous because the subject is indefinite. Indefinite subjects in Icelandic may potentially occupy a low subject position, and so it is possible that the verb has moved no higher than Tense in the sentence above. (The lower position is below Spec(TP), possibly Spec(vP), but the exact position is unclear; see Rögnvaldsson 1984a,b for the original description of the phenomenon and Thráinsson (2007: 26-27) for an overview of more recent approaches to this issue).

yesterday laid **REFL** mother early
"Yesterday mother went to bed early."
(Josefsson 1992: 67)

This shared feature among the scrambling and object shift languages is another indication that scrambling and object shift are in fact a unitary phenomenon, and modern Yiddish scrambling appears to be an historical stage between the two systems. Like Swedish object shift, the reconstruction effect can only be seen in a type of scrambling that is parasitic on verb-movement: the object can only scramble past subject in Spec(TP) if the verb moves to C, as in Swedish. However, because of Yiddish's continuing OV nature, the phenomenon is not parasitic on movement of the lexical verb in a clause, but only on the finite verb (potentially a finite auxiliary, as in the examples above). In this way, the change we have observed in the position of Tense in Yiddish and its effect on object scrambling shows how a German-style scrambling system can evolve directly into a Swedish-style one through changes in the language's functional structure.

3.5 Summary and Conclusions

The results for historical Yiddish are simple to describe: when Tense-medial clauses begin to appear, scrambling to clause-level (higher than Tense) is not allowed in those clauses. Using a parsed corpus of early Yiddish (Santorini 1997/2008), the chapter showed that the variation in East Yiddish accompanying the gradual change in the

position of Tense shows a striking gap in the set of logically possible combinations of scrambling and the position of the tensed verb (Tense-initial plus clause-adjoined scrambling). Additionally, clauses with objects to the left of the finite verb decline at the same rate as clauses that are unambiguously Tense-final on other criteria (slopes of -.0142 and -.0147 respectively), showing that they are the same phenomenon, and no scrambling across the verb took place. The diachronic data provides a type of information that the synchronic typology cannot: once a phrase structure change begins, the resulting variation within a single speech community and even within a single speaker immediately shows the effect of the GHC on scrambling. Note that this is very similar to the argument from Old English in Pintzuk (1991) and Kiparsky (1996) for the universal non-existence of SVOI orders. In a sense, this type of data is more conclusive than a comprehensive synchronic typological study ever could be.

Arguments from synchronic typology ultimately rely on the field's knowledge of human language varieties (which is considerably less than exhaustive), as well as some basic consensus about how the syntax of all of the languages considered should be analyzed. In a study like the present one, on the other hand, there is a clearly defined and highly restricted domain of inquiry. There is one community of speakers involved, the community is already well-studied, and the variation under investigation is the result of change in a single syntactic parameter (and there is some consensus about the structure of the rest of the system surrounding that parameter). Most importantly, the structural variation accompanying the parameter change is represented in the minds of individual speakers as well as at the community level (Santorini 1992). Thus, if a particular

interaction between the parameter variation and the rest of the system is not found (i.e. Tense-medial with scrambling above Tense), it can be counted a true gap and must be the result of some non-accidental force.

Chapter 4

Early Modern English Scrambling/Object Shift and the V-to-T Movement Parameter[26]

4.1 Introduction

Although it has been noticed from time to time in the literature that unstressed pronouns in English show some syntactic behavior that is not shared by their full DP counterparts, these differences have not yet been studied and accounted for in a systematic way.[*] This is not particularly surprising, as the special behavior of English pronouns is confined to only a few contexts, and so very few clues are provided to the analyst about the underlying processes involved (in contrast to, e.g., pronoun scrambling in German or Dutch). But this is only the case if we confine ourselves to data from the modern language. It is possible to learn a great deal more, however, by taking a broader, historical perspective of the phenomenon. With the help of corpus data from earlier stages of English, I will argue that the exceptional syntax of modern English object pronouns actually results from their status as "weak pronouns", in the sense of

[26] This chapter is a revised and expanded version of previously published work in Wallenberg (2007).

Cardinaletti & Starke (1999), and their ability to undergo scrambling, resulting in a pattern like "Object Shift" of the Scandinavian type (Holmberg 1986, Holmberg 1999, Thráinsson 2001 *inter alia*). However, scrambling/object shift is severely restricted in modern English by a combination of Holmberg's Generalization/GHC and an independent, but intersecting, phenomenon: the loss of movement of main verbs to Tense in the overt syntax. In this way, English is not the odd language out in Germanic, but rather patterns with the rest of the family in having a leftward movement process that applies to its unstressed object pronouns. It just so happens that this object shift is not very easy to see, because it has been obscured by the loss of V-to-T movement. In fact, when one controls for this change in verb movement, it becomes apparent that English object shift has been remarkably stable over time during the history of the language.

In the first section below I present some examples of the special syntax of weak object pronouns in modern English, and provide an analysis in terms of scrambing/object shift and phrasal adjunction. Section 3 discusses the history of object shift in English, which supplies further evidence in favor of the object shift analysis of modern English, using two parsed corpora of early English: the Penn-Helsinki Parsed Corpus of Middle English ("PPCME2", Kroch & Taylor 2000b) and the Penn-Helsinki Parsed Corpus of Early Modern English ("PPCEME", Kroch et al 2004). Finally, I offer some conclusions in Section 4.

4.2 Weak pronouns in modern English

In this section, I argue that the facts about the syntax of English unstressed object pronouns are best understood by taking them to be *deficient*, or *weak pronouns* in the sense of Cardinaletti & Starke (1999): they are phonologically weak, leaning leftward on a (verb) host if possible, and they obligatorily move from their base (theta) positions. As Cardinaletti & Starke (1999: 170) point out, however, weak pronouns are not true clitics in that they are not heads. They are still phrasal, and as such, they can adjoin to maximal projections as "phrasal affixes" (cf. Klavans 1985). I suggest that the Scandinavian "object shift" phenomenon (Holmberg 1986) is a subtype of weak pronoun movement with phrasal affixation, and that the behavior of English pronouns is also an instance of object shift. As such, the leftward movement of English pronouns is expectedly constrained by the position of their theta-assigning verb, according to the Generalized Holmberg Constraint (and Holmberg's original Generalization, in this particular case): they cannot move leftward to any position past the final landing site of the main verb in the clause. The landing sites of weak pronoun objects are also sites that are in the extended projection of the verb, e.g. TP, vP; in other words, the sites of phrasal affixation should form a natural class, as in the "landing sites" theory of Baltin (1982). In Scandinavian, it is easy to observe the leftward movement of pronoun objects because the finite verb moves leftward to T (in all clauses in Icelandic) or to C (in matrix clauses in the mainland Scandinavian languages), and so object pronouns frequently shift across negation, adverbs, and other elements inside the vP. In English, on the other hand, non-auxiliary verbs do not move any farther left than little-*v*, even if they are finite. This

makes object shift even more restricted in English than it is in Scandinavian, and much harder to observe. However, there are still a few constructions which hint at the presence of object shift and the existence of weak pronouns in modern English.

The most obvious of these constructions is the well-known verb-particle alternation, which is illustrated in (1) and (2) below:

(1) The poodle ate my supper up.

(2) The poodle ate up my supper.

Unstressed pronouns, unlike full DP objects, do not participate in this alternation, as the two sentences below demonstrate.

(3) Irene threw it out.

(4) *Irene threw out it.

The order in (2) and (4) is not available to pronouns unless they are contrastively stressed, as in (5), and it is not available at all to pronoun objects that cannot be stressed, such as *it* in (4) above, or the weak form of the 3pl object pronoun, *'em*, as in (6) and (7) below.

(5) I didn't pick up HIM, I picked up HER!

(6) I threw 'em out.

(7) *I threw out 'em

Note that the 3pl pronoun also clearly has a distinct phonological form for its weak version which is not the product of general phonological rules (though the same might be true of *'im* and *'er* for *him* and *her*, which are also restricted as in 7); having a special form associated with a special syntactic position is characteristic of the weak pronouns discussed in Cardinaletti & Starke (1999), as well as the Swedish and Norwegian object-shifting pronouns in Hellan & Platzack (1995).

Following Svenonius (1996; see also den Dikken 1995 for a similar proposal), I will assume that the particle and object form a small clause complement to the verb, with the particle as a PP predicate of the small clause, as in (8).

(8) The poodle ate $[_{SC}\ [_{DP}$ my supper$]\ [_{PP}$ up$]]$

However, the particle is simultaneously minimal and maximal (P and PP), and is able to undergo a short movement to a head (called Pred in Svenonius 1996) at the left of the small clause, as in (9):[27]

(9) The poodle ate $[_{PredP}$ up$_i$-Pred $[_{SC}\ [_{DP}$ my supper$]\ [_{PP/P}$ t$_i$ $]]]$

This derives the two verb-particle orders with DPs. Weak object pronouns, on the other hand, do not alternate with particles, because they must move from their base positions (as in Cardinaletti & Starke 1999) and left-adjoin to a maximal projection

[27] Whether or not this movement of the particle is truly head-movement is not crucial to the analysis presented here. The key point is that the particle undergoes a short movement to the left periphery of the small clause. Similarly, it is not necessarily the case that the Pred head is present both in the order in (8) and in the order in (9), though Svenonius does asssume this (i.e., it may actually be absent in the type of sentence in 8).

associated with their theta-assigning verb. At this point, it's necessary for me to say a bit more about the structure of the verb phrase. As in Embick & Noyer (2001) and other proposals in the Distributed Morphology framework, I accept the "vP" hypothesis of Kratzer (1996) and interpret within DM to mean that the traditional "VP" is composed of a vP, headed by a functional verbalizing head *v*, which takes as its complement a RootP (this projection corresponds to the "VP" in many Minimalist proposals). The RootP is headed by the lexical "verb", which is in fact category-neutral, and must therefore move out of the RootP by head movement to incorporate with the *v* head and become verbalized. Thus we can more fully articulate the structure in (9) as (10) below:

(10) The poodle $[_{vP}$ ate$_j$-v $[_{RootP}$ $[_{Root'}$ t_j $[_{PredP}$ up$_i$-Pred $[_{SC}$ $[_{DP}$ my supper] $[_{PP/P}$ t_i]]]]]]

It is now clear what the effect of object-shifting a weak object pronoun will be for this structure. Because a weak pronoun, like *it*, cannot remain in its base position in the small clause, it must move to left-adjoin to some projection of the verb. However, it cannot cross the final landing site of the verb at *v*, as per Holmberg's Generalization. Therefore the only landing site available to the object pronoun is RootP. For this reason, the unstressed, unfocussed pronouns will always appear to the left of the particle, whether or not the particle movement has applied. This is shown in the two possible structures for (11), in (12) and (13) below.

(11) The poodle ate it up.

(12) The poodle $[_{vP}$ ate$_j$-v $[_{RootP}$ it$_k$ $[_{RootP}$ $[_{Root'}$ t_j $[_{PredP}$ Pred $[_{SC}$ $[_{DP}$ t_k] $[_{PP/P}$ up]]]]]]]

(13) The poodle [$_{vP}$ ate$_j$-v [$_{RootP}$ it$_k$ [$_{RootP}$ [$_{Root'}$ t$_j$ [$_{PredP}$ up$_i$-Pred [$_{SC}$ [$_{DP}$ t$_k$] [$_{PP/P}$ t$_i$]]]]]]]

This adjunction to the RootP also licenses the appearance of weak object pronouns in the simple case, sentences without particles. That is, when a verb takes a single object, it is possible for that object to be a weak pronoun, as in (14) below, because of the presence of the vP and RootP projections; the result does not become ineffable. Of course, the leftward movement is string-vacuous in this case:

(14) The poodle ate {his supper / my socks / the woman / the man}.

(15) The poodle ate {it / 'em / 'er / 'im }.

(16) The poodle [$_{vP}$ ate$_j$-v [$_{RootP}$ 'em$_k$ [$_{RootP}$ t$_j$ [$_{DP}$ t$_k$]]]]

And in (15), the weak pronouns avoid the sentence-final intonational peak that falls on the object in (14) because they are leftward-leaning phonologically, and can form a prosodic word with the verb, thereby forcing the intonational boundary tone to fall on the verb (i.e. the first syllable of the verb+pronoun phonological complex). Again, this is similar to the Swedish and Norwegian object-shifting pronouns, which are syllabified with the verb when they immediately follow it (Hellan & Platzack 1995). Additionally, a nice result of this analysis of verb-particle constructions is that it formalizes the astute observation in Rögnvaldsson (1982) that English particle shift and Scandinavian object shift appear to be parallel and are probably related.

With the analysis above, it is now also possible to understand a peculiar aspect of the English double-object construction. Larson's classic study of the double-object

construction (1988: 364) noted that it is not possible for the direct object of a double-object sentence to be an unstressed pronoun if the indirect object is a full DP, as in (17).

(17) * John gave the boy it.

It is, however, possible for the indirect object to be a pronoun when there is full DP direct object, as in (18), or for both objects to be pronouns, as in (19):

(18) John gave him the book.

(19) John gave him it.

Note that this effect cannot be analyzed as purely the result of a phonological requirement for the unstressed pronoun to be adjacent to the verb. In (20), the sentence is grammatical even though the unstressed pronoun is not adjacent to the verb, by virtue of the fact that the pronoun is inside a QP.[28]

(20) John had no money left, because he'd given the boy it all.

It was suggested to me by Josef Bayer (p.c.) that the grammaticality of (20) does not necessarily argue against an analysis under which the effect in (17) is due to a PF constraint on *it*, because *it* in (20) could potentially be exempt from this constraint by being phonologically parsed with the QP, rather than with the preceding material. Under this story, (17) would be ungrammatical because English weak pronouns must occur

[28] It is true, however, that *it* still sounds most natural when immediately adjacent to the verb and parsed with it phonologically.

adjacent to the verb (be enclitic on the verb) if they occur in the same phonological phrase as the verb, but not otherwise. However, if we also consider the case of the weak 3pl pronoun *'em*, it becomes clear that this is not the entire story. The pronoun *'em* is obligatorily unstressed, like *it* is, but unlike *it*, *'em* truly does have a PF requirement that it be enclitic on the verb. Because of this requirement of phonological enclisis, *'em* cannot occur in sentences like (20):

(21) * John had no M&Ms left, because he'd given the boy 'em all.

Similarly, *'em* cannot occur when it is separated from the verb by a particle, even if it is within a QP, as in (23).

(Who picked up all the marbles?)

(22) John picked 'em all up.

(23) * John picked up 'em all. (Even with stress on *all*.)

The position of *'em* in (23) is not the preferred prosodic position for *it* either, but *it* is much better in that position than *'em*, particularly if one is careful to stress *all* in sentences like (25):

(Who cleaned up the spilled milk?)

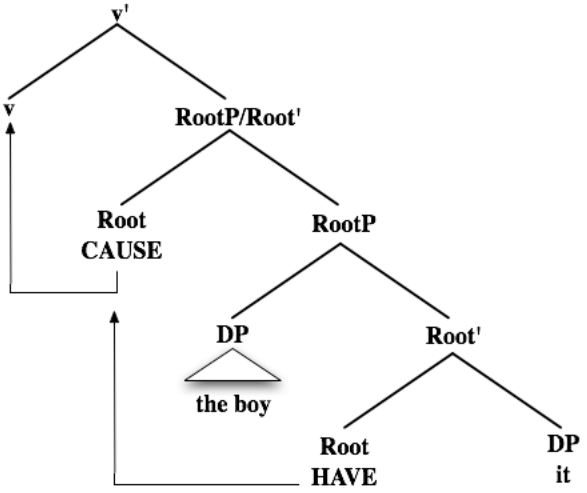
(24) John cleaned it all up.

(25) John cleaned up it all.

Therefore, it is clear that phonologically parsing a pronoun with the QP is not sufficient to escape a PF constraint on enclisis, and the explanation for the contrast between (17) and (20) must lie elsewhere. The effect in (17) must be the result of a syntactic property the pronoun has when it occurs alone, rather than because of some surface phonological constraint. This syntactic property is the requirement that weak pronouns leave their theta positions.

If we adopt a Larsonian VP-shell analysis for the double-object construction of the type proposed in Harley (2002), and then transpose it into the theory of the vP that I sketched out above, then the structure of the vP in (17) would be as below.

(26)



In this type of shell analysis, ditransitives such as *give* are semantically decomposed in CAUSE and HAVE predicates, and so there is a recursive RootP here for

give. CAUSE and HAVE then compose in the syntax via head-movement, and then become verbalized by moving again to incorporate with *v*. The weak pronoun *it* must leave its base position to adjoin to one of the RootPs. However, it cannot move to the lower RootP to derive **gave it the boy* because of an independent constraint which prevents it from moving across the direct object *the boy*. It is beyond the scope of this dissertation to provide a full discussion of the nature of this constraint, but Bobaljik (2005: 123-4) has shown that the constraint is independent of the other workings of object shift. He points out that languages can differ on whether or not this constraint applies, since some Swedish and Norwegian speakers can object shift a direct object over the indirect, as in (27), while Danish does not.

(27) Jag gav den inte honom.
I gave it not him.
"I didn't give it to him."

(28) ?Han gav den inte henne.
I gave it not her.
"I didn't give it to her."

(29) Han gav'en inte'na.
I gave it not her.
"I didn't give it to her."
(examples 28, 29 are Swedish, from Hellan & Platzack 1995: 58)

Furthermore, Norwegian also allows the passivization of a direct object over an indirect ("symmetric passives"), while Danish does not, and so this constraint on movement in double-object constructions is clearly an independent principle of the grammar. An example of this type of passive is cited in Holmberg & Platzack (1995: 215), and

discussed again in Bobaljik (2005). Holmberg & Platzack (1995: 215) report that examples such as (31) are allowed in Norwegian and "somewhat marginally" in Swedish alongside (30), but are ungrammatical in Danish:

(30) Jens blev givet bogen.
Jens was given book-the
"Jens was given the book."

(31) Bogen blev givet Jens.
book-the was given Jens.

This correlation between symmetric passives and symmetric pronoun movement is also well established outside of Scandinavian: it is a known typological correlation in the Bantu family. There the type of pronoun movement at issue is the incorporation of pronouns into the finite verb, and the languages differ as to whether they can incorporate a direct object pronoun in a double object construction where an indirect object has been left unincorporated (e.g. the indirect object is a full DP, and so does not incorporate). Bresnan & Moshi (1990: 150-151) show that Kichaga can incorporate either or both objects in the Bantu double-object applicative construction (both of which pervert the base order of the two objects), whereas Chichewa cannot. Not surprisingly, Kichaga also allows passivization of the lower object over the higher (as well as other symmetric constructions), while Chichewa does not (Bresnan & Moshi 1990: 150).

Interestingly, Early Modern (British) English did allow shifting of a direct object over an indirect, as in (23), unlike modern American English; there are 6 such examples in the PPCEME, and no examples parallel to (17), with an in situ direct object pronoun.

Not surprisingly, Early Modern English also allowed symmetrical passives, as in (24), whereas this is not possible in modern American English; 14/129, or 10.9%, of the ditransitive passives in the PPCEME are of this type.

(32) I think he will carry this island home in his pocket, and give it his son for an apple.
(Shakespeare, *The Tempest*, II, i, 92-93)

(33) ...and those were sent a Friend of mine for a present.
(Elizabeth Oxinden, *The Oxinden and Peyton letters*,
EOXINDEN-1 660-E3-H,308.10 in the PPCEME, date: 1642-1670)

These passives are not possible in modern American English, and so the two dialects of English fall neatly into the typology seen in the Scandinavian languages and in Bantu.

Returning to the structure in (26), for languages like Early Modern English or Swedish, the weak pronoun in the object position can indeed move to left-adjoin to the lower RootP. However, for modern English, it can only move to the lower RootP if the indirect object has also moved, to adjoin to the higher RootP. This is the derivation of the *John gave him it* in (19). And, trivially, if the higher object is a weak pronoun and the lower one is not, it moves to adjoin to the higher RootP, deriving the pattern in (18). The sentence in (32) is also a particularly good example of this context for object shift, because there is an auxiliary in the clause, *will*, and so *give* cannot have moved to Tense to license the shifting of *it* over *his son*. This shows unambiguously that object shift operated in English even in the absence of verb-movement to T.

Finally, there is another construction that indicates the presence of two sets of object pronouns in English with different syntax, one strong, the other weak. It is a very

colloquial construction with an ancestry dating back to the Old English case-marking system, which I will refer to as the *dative-benefactive construction*. It occurs in sentences like (34), which involve a non-reflexive object pronoun that is nevertheless coreferential with the subject of the clause.[29] (34) corresponds to the standard English *I'm going to get myself a beer*.

(34) I_i'm gonna get me_i a beer.

This construction is interesting in that the object pronoun is restricted to be of the strong, stressed variety. The weak pronouns are not available in this construction under the interpretation in which the object pronoun is coreferential with the subject, as one can see in the contrasts below.[30]

(35) $They_i$'re gonna get $them_i$ a new car.

(36) * $They_i$'re gonna get'em_i a new car.

(37) $They_i$'re gonna get'em_j a new car.

Similarly:[31]

(38) He_i needs to get him_i a new car.

(39) * He_i needs to get'im_i a new car.

---

[29] This apparent binding violation probably shows that this construction is a retention from a time in the history of the language before reflexive pronouns developed.

[30] Several native speakers of American English agreed on the judgments above, and they also represent my own judgments. This construction is nonstandard and stigmatized in the U.S., so it is not particularly easy to get clear judgments on it. However, all of the speakers I have questioned (5/5, not including myself, and including one Texan speaker from a core area of this construction) who accept (25) and (27) also agree with the remainder of the judgments.

[31] Thanks to Marjorie Pak for pointing out the parallel facts with regard to *him* vs. *'im* in (29)-(31).

(40) He_i needs to get'im_j a new car.

The fact that this construction requires the strong versions of the English pronouns is further confirmation that their syntactic distribution is not precisely the same as that of the weak pronouns. And under an interpretation other than the dative-benefactive, there is nothing preventing the weak pronouns from appearing in what is otherwise the same syntactic position, as in (37) and (40). Zwart (1996a: 584) makes the same argument for Dutch on the basis of a similar type of reflexive construction, which is restricted to the weak, rather than the strong set of object pronouns. An example is in (41) below:

(41) Jij schaamt {je / *jou}.
"You're ashamed"

A related construction, which is even more similar to the English dative-benefactive than the Dutch one, exists in a variety of Bavarian German (Josef Bayer p.c.).[32] In this construction, a seemingly non-reflexive pronoun can be used with reflexive reference, as in (44), instead of the reflexive pronoun in (43).

(42) Er hot-se am Koopf kratzt
he has-REF at-the head scratched
"He_i scratched his_i head"

(43) Er hot eam am Koopf kratzt
he has him at-the head scratched
"He_i scratched his_i head"

[32] Thanks to Josef Bayer for writing to me about these facts. He described the dialect/style in question as "parochial", but reports clear judgments on the sentences.

Like the English dative-benefactive, reflexive reference is possible only when the strong pronouns are used. Bayer (p.c.) states that, in examples such as (44) below, "when I use a clitic (or weak pronoun), I am forced to a disjoined interpretation."

(44) Er hot-n am Koopf kratzt
he has-him at-the head scratched
"He$_i$ scratched his$_{j/*i}$ head"

Apparently, colloquial English's dative-benefactive construction and its sensitivity to the weak/strong pronoun distinction has plenty of good company in the rest of Germanic.

4.3 Evidence from historical English: what doesn't change, doesn't change.

The analysis above becomes more plausible in light of diachronic data, which shows object shift to be highly stable over time. Verb movement, on the other hand, becomes more restricted over time in English, as overt movement of the verb to T is gradually lost and *do*-support enters the language, beginning around the year 1400 (Kroch 1989). Consequentially, examples of object shift become more and more rare over time, as the leftward movement of the object pronoun is bounded by the position of the verb to its left (Holmberg 1986).[33] However, it is possible to show that the object shift

[33] Of course, this discussion only refers to VO clauses, in which the verb to the immediate left of the object will always be a non-auxiliary verb, the which assigns a theta-role to the object. However, as Kroch &

rule is statistically stable even as verb movement is in flux, by looking at clauses which contain some element that is both diagnostic for verb movement and for the leftward movement of object pronouns. Once the verb-movement parameter is controlled for, the statistical study shows that object shift never really changed in English, and minus any other interfering factor, it should still be present in the language today.

Middle English and Early Modern English still showed a large amount of V-to-T movement with finite main verbs, and so it is not difficult to find unambiguous examples of object shift. Although leftward shifting of the pronominal object was still bound by the position of its licensing verb, finite main verbs appeared in T, and so the object was able to shift leftward past sentential negation and adverbs, in the modern Scandinavian way; compare the modern Icelandic sentence in (45) to the Middle English sentence in (46) and the Early Modern English sentence in (47) (the shifted object pronouns are in boldface):

(45) …að þeir lásu hana ekki.
that they read her not
"that they didn't read it" (Hellan & Platzack 1995: 53)

(46) Whi telles tu mi rihtuisnes wid þi muz, and dos **it** noht?
Why tell you me righteousness with the mouth, and does it not
"Why do you talk to me about righteousness but don't do it?"
(*The Northern Prose Rule of St. Benet*, CMBENRUL,5.152 in P PCME2, date: 1425)

(47) but if thy conscience condemne **thee** not, I thinke thy sinne one of the least sinnes;

Taylor (2000) have shown, Middle English contained a mixture of underlyingly OV and VO clauses. It is beyond the scope of this paper to discuss the interaction between the change from OV to VO in English and object shift, but this will be the topic of future publications.

(Henry Smith's *A preparative to mariage*, SMITH-E2-H,E7R.311 in PPCEME, date: 1591)

In all three sentences, the verb has moved to T, and the object has also moved from its base position, appearing to the left of negation. Note also the late date of the example in (47).

Earlier stages of English also allowed another type of object shift, in which a verb has moved higher in the structure to C, and an object pronoun has consequently shifted further leftward to a position higher than T. This V-to-C movement lands the verb in a position that is higher than, and so to the left of, the subject in Spec(TP), and also licenses the further shifting of an object pronoun past the subject, as shown in (48) and (49):

(48) ne set **me** neauer na þing se luðere ne se sare.
NEG oppressed me never no thing so painful nor so sore
"Nothing ever oppressed me so painfully or so sorely."
(*St. Juliana*, CMJULIA,112.280 in PPCME2, date: 1225)

(49) Then answered **them** the Pharisees, Are ye also deceiued?
(*The Holy Bible Authorized Version*, AUTHNEW-E2-H,VII,40J.968 in PPCEME date: 1611)

Both (48) and (49) involve well-established V-to-C triggers in historical English (cf. Kroch & Taylor 1995 and Kemenade 1987: 138-9): (48) has a negated verb with prefixed negation (which was common in Old English and found in the more conservative ME dialects), and (49) contains the trigger "Then" in Spec(CP). In (36), the object pronoun *me* has moved not only past the subject *na þing*, but also past the clausally-adjoined adverb *neauer*. Incidentally, this type of example also provides an

argument that object-shifted pronouns are adjoined to maximal projections: the only possible position for the pronoun in (36) is left-adjoined to the TP just as *neauer* is, under the standard assumption that it is not possible for a pronoun to be right-adjoined to a head via head-movement (cf. Kayne 1991, Kayne 1994). The pronoun *them* in (49) has similarly shifted to a position preceding the subject *The Pharisees*.[34] And again, note the late date of the example in (49), contrasting with the early date of the example in (48). The latter also occurs in a text from the conservative Middle English dialect area in the Southeast of England (cf. Kroch & Taylor 1995, 2000). Finally, this analysis of these examples, as involving a type of object shift under V-to-C movement, is confirmed by the fact that this same phenomenon is found in modern Swedish; it is called "long-object-shift" in Josefsson (1992) and Hellan & Platzack (1995). Compare the English examples above to the Swedish one below:

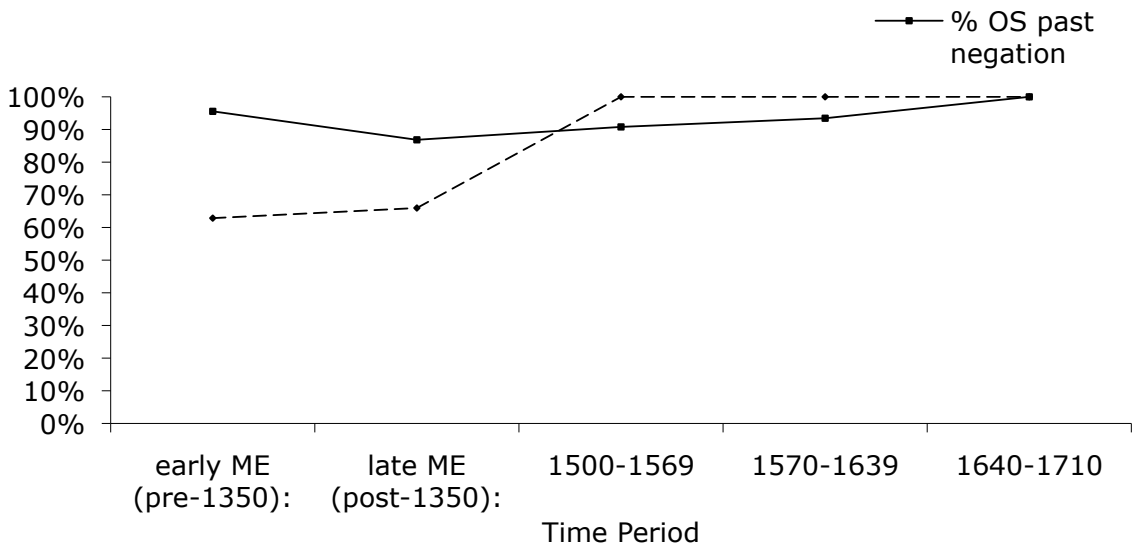
(50) Igår kammade **sig** Erik inte på hela dagen.
Yesterday combed REFL Eric not on whole day-the
(Hellan & Platzack 1995: 58)

The examples above in (45)-(50) serve to illustrate the two diagnostic contexts, sentential negation and the position of the subject, which allow us to estimate the frequency of object shift in the various stages of English. When the negation context is restricted to clauses in which the finite verb appears to the left of negation, it becomes diagnostic for V-to-T movement as well. Similarly, the long-object-shift context only

[34] Note also that this example and other similar examples considered in this study are unlikely to be derived by postposition of the subject: subjects like *the Pharisees* in (49) are not particularly heavy, and none of the examples considered here are unaccusatives (by definition, as the verbs all have objects), which are the usual environments for subject postposition throughout the history of English (Warner 2007).

includes clauses in which a finite main verb has inverted with the subject, and so the verb must have undergone V-to-T-to-C movement. In this way, it is possible to factor out the effect of the historical change in verb-movement, holding that parameter constant across the time variable. The chart below plots the frequency of object shift across negation (i.e., the string **finite-V > object-pronoun > neg** vs. the **finite-V > neg > object-pronoun** order), as well as the frequency of long-object-shift across a full DP subject, out of all the relevant clauses containing inversion of the subject and a finite verb (**V > obj-pron > sbj** vs. **V > sbj > obj-pron**). The time period over which the frequencies are plotted begins in the 12th century, in Early Middle English, and ends at the end of the Early Modern English period in the 18th century. As the chart below shows, the frequencies are essentially stable over more than five centuries, and in both object shift contexts, the shifted order is consistently preferred over the in situ order.

Figure 1.



If anything has changed at all over this time period, it is that object shift actually becomes somewhat more frequent in the Early Modern English period, after 1500, though it is also possible that this effect is merely an artifact of the small number of examples that the later frequencies are based on. As one can see in the tables below, the N for both object shift contexts drops off as one goes forward in time, even though the frequency of object shift remains nearly constant. This shows that the syntax of object shift in English remains unchanged even as the loss of V-to-T movement goes to completion. And, of course, the loss of V-to-T movement bleeds movement of the main verb to C as well, and so the long-object-shift context also eventually disappears (to be replaced by *do*-support, as shown in Kroch 1989 and Kroch & Han 2000).

Table 5.
Object-Shift of Object Pronouns Past NEGation in Middle English and Early Modern English, matrix and subordinate clauses containing finite main verbs

	V > pron-Obj > NEG	V > NEG > pron-Obj	TOTAL N	% object shift
early ME (pre-1350)	43	2	45	95.56%
Late ME (post-1350)	165	25	190	86.84%
1500-1569	69	7	76	90.79%
1570-1639	114	8	122	93.44%
1640-1710	15	0	15	100.00%

Table 6.
Long-Object-Shift of Object Pronouns past DP Subjects in Middle English and Early Modern English, matrix and subordinate clauses containing finite main verbs

	V > pron-Obj > Sbj	V > Sbj > pron-Obj	TOTAL N	% long-OS
early ME (pre-1350)	22	13	35	62.86%
Late ME (post-1350)	31	16	47	65.96%
1500-1569	12	0	12	100.00%
1570-1639	3	0	3	100.00%
1640-1710	5	0	5	100.00%

The ultimate effect of the loss of V-to-T movement on object shift is to cause object pronouns to be trapped below *v* by the Holmberg's Generalization constraint, as their theta-assigning main verbs no longer move any farther than *v* in the modern language. But the diachronic stability of object shift, even into the late period when V-to-T movement had already been mostly lost, suggests that the syntax of object shift in English never really changed at all.

Finally, this interpretation of the diachronic data makes a very clear prediction about modern varieties of English: if object shift never changed, modulo the effects of verb movement, then any variety of English which still allows overt verb movement should also allow object shift. This prediction is borne out in a dialect of Belfast English, described as "Dialect B" in Henry (1995: Chapt. 3). As Henry (1995) argues, Dialect B of Belfast English requires overt movement of the lexical verb to C in positive imperatives. This results in subject-verb inversion, as in (51) and (52):

(51) Read you that book.

(52) Do you your best.
(Henry 1995: 55)

As predicted, this verb movement also triggers the kind of long-object-shift that's found in Early English and modern Swedish: the verb movement licenses movement of weak object pronouns to the left of the subject. This is shown by the position of *them* in (54) and *it* in (55).

(53) Eat them you now. (Henry 1995: 65)

(54) Read it everybody quickly. (Henry 1995: 71)

The position to the left of the subject is, of course, not available to full DP objects in Belfast English:

(55) a. * Give the book you to the teacher.
b. Give you the book to the teacher. (Henry 1995: 58)

Nor is this position available to stressed or coordinated pronouns, both of which must involve strong pronouns according to Cardinaletti & Starke (1999).

(56) a. Put them you in the living room.
b. Put you them in the living room.

(57) a. * Put them and me you in the living room.
b. Put you them and me in the living room.

(58) a. * Put THEM you in the living room.
b. Put you THEM in the living room.

(59) a. * Put THEMUNS you in the living room.
b. Put you THEMUNS in the living room.
(Henry 1995: 71)

These facts provide further evidence that modern English varieties have different weak and strong pronoun sets, and that it is the weak set which must undergo leftward movement. This obligatory nature of the leftward movement can also be seen clearly in Belfast English. While shifting a weak pronoun to the left of the subject in Belfast English imperatives is optional, as in (56) above, the position of sentential adverbs shows

that some movement of the object pronoun has still taken place even in examples where the pronoun surfaces to the right of the subject. Even if the object pronoun remains lower than the subject, it must still shift to a position above sentential adverbs (as well as vP-adjoined adverbs, of course) in this dialect (Henry 1995: 72).[35] This is apparent from the contrasts below:

(60) a. Read it always you to me.
b. Read you it always to me.
c. * Read always it you to me.
d. * Read you always it to me.

The data from Belfast English's Dialect B are all the more surprising in light of the fact that Belfast English, like Irish-English (and other varities) more generally, has largely lost V-to-T and V-to-C movement of lexical verbs. Belfast English's Dialect B maintains robust V-to-C movement of lexical verbs only in the case of positive imperatives, and so speakers of this dialect can only shift object pronouns past the subject or past adverbs in this construction, as Holmberg's Generalization predicts. In this way, just as object shift is more visible in different stages of English, because it depends on verb-movement, object shift is more visible in this construction within Belfast English than it is in other constructions, because the conditions on verb-movement differ. Belfast English thus presents another striking illustration of how object shift can remain robust even as it is partially hidden by changes in verb-movement. But if the verb is allowed to move past little-*v*, even if the movement occurs only in a small corner of the grammar,

[35] The question still remains as to why dialects of English and the Scandinavian languages differ as to how far weak pronouns can or must move. I leave this point of fine dialect variation for discussion in later sections.

weak pronouns are allowed to shift to the same higher positions that the historical language allowed.

4.4 Conclusion

This chapter has argued that there are two sets of pronouns in English: one stressed set, and one deficient or weak set. This study has further shown that the two are distinguished by different syntax, and that the exceptional syntax of the weak pronouns may be observed throughout the recorded history of the language. This diachronic continuity, as well as the syntactic patterns themselves, shows English to be more alike to the rest of the Germanic family than different from it in its pronominal system and its constraints on pronoun scrambling. English pronouns undergo scrambling/object shift just as Scandinavian pronouns do, modulo the effect of a different verb-movement system. As finite-verb movmenet to T declines over time in the Early Modern English period, the effect of the GHC makes itself known and begins to severely restrict the ability of English weak pronouns to scramble leftward. This shows the diachronic stability of the GHC, which is expected if it is a component of UG.

Chapter 5

Scrambling and the Antisymmetry Hypothesis

"C'est que la symétrie, c'est l'ennui" –Victor Hugo, *Les Miserables*

5.1 Introduction

This chapter takes a step back from the empirical issues surrounding the GHC and focuses on the exact statement of the GHC and its place in the architecture of the grammar. In making the definition of the GHC as precise and general as possible, it will become clear that the cross-linguistic behavior of scrambling bears on the much larger theoretical issue of the antisymmetry hypothesis (beginning with Kayne 1994). In the following chapter, I will argue that the empirical result of the GHC actually allows us to decide between the antisymmetric and classical "head parameter" approach to representing the distinction between "head-final" and "head-initial" languages. Unlike many components of the syntax, which can be stated equally well under a classical or antisymmetric-style phrase structure (for which the two approaches function as mere

notational variants), the GHC is actually not as easily statable under classical symmetric assumptions. Only an antisymmetric approach to head-finality allows the GHC to be reformulated in purely hierarchical terms, and as a general property of the adjunction operation ("Conservation of C-Command"). I will show that the constraints on scrambling constitute a subcase of this general constraint on adjunction. Additionally, building on the proposals in Biberauer (2003a,b) and Biberauer & Roberts (2005), I will show how the combination of an antisymmetric approach to Germanic head-final languages and the scrambling operation also provides a straightforward account of the various patterns found in the West Germanic verb-raising and verb-projection-raising constructions (cf. Evers 1975, Zaenen 1979, Haegeman & van Riemsdijk 1986, Kroch & Santorini 1991, Wurmbrand 2004, 2005).

5.2 Adjunction, Bare Phrase Structure, and Antisymmetry

Now that the empirical validity of the GHC is clear, it is appropriate to return to the question of its theoretical status. In particular, is the constraint compatible with standard assumptions about the grammatical architecture? And is it stated in the most general form possible (and therefore, scientifically most interesting form)? I suggest that the answer to both questions is "no", as things stand. First of all, it has been standardly assumed for some time (probably since Chomsky 1986) that heads and dependents do not enter the derivation with some particular precedence relation defined between them;

linearization takes place at some later stage of the derivation,[36] and so the implication that "left" and "right" are primitive notions in the phrase structure must be removed from the statement of the GHC if it is to be compatible with the standard assumption that most (if not all) of the syntactic computation refers to hierarchy alone. Secondly, it would be ideal from a purely methodological view to reduce the GHC to a constraint on either linear precedence, or a constraint on hierarchical structure, without having to refer to both notions. However, it is not immediately obvious how to dispense with either part of the definition. According to the structures I have given for, e.g., German on the one hand and Yiddish on the other, the initial or final placement of the Tense node is crucial for determining scrambling options, so it is not possible to remove precedence from the GHC. One might then attempt to state the GHC solely on precedence, but then the fact that elements can scramble past AdvPs, for instance, would be impossible to state: an AdvP in German or Icelandic contains a head to the left of a potentially scrambling object, but the Adv head crucially does not c-command the verb's theta positions.[37] And unfortunately, under the classical assumptions about phrase structure that I have assumed thus far, there is no obvious way to derive precedence from dominance or vice-versa.

---

[36] Whether linearization takes place as the very last step, which is implicit in the Distributed Morphology literature (see Embick & Noyer 2001 for an example) or whether pieces of the derivation are linearized before the entire sentence is finished, as in phase theory. See also Fox & Pesetsky (2005) for a type of compromise between the two views, in which whole chunks of the derivation are not linearized before continuing the derivation, but precedence relations are added as the syntactic computation proceeds and conjoined in a final statement at the conclusion of the derivation.

[37] Stating the GHC in terms of precedence alone would presumably be equivalent to proposing that scrambling is PF movement, as in Holmberg (1999). Note that the proposal Holmberg (1999) has no natural way to state the fact that scrambling can proceed past adjuncts, and in some languages past arguments (e.g. German, Norwegian – cf. Bobaljik 2005: 123, and Early Modern English as in Chapt. 4 above).

Fortunately however, the shortcomings of the GHC are entirely due to its being couched within a classical, X-bar-theoretical phrase structure (in the tradition of Chomsky 1986), in which languages differ by settings of a "headedness parameter"; in this type of framework, OV and VO languages, for instance, are identical in hierarchical structure and so their phrase structures are mirror images. On the other hand, under a version of the antisymmetry hypothesis (Kayne 1994 and much subsequent work) along the lines of Biberauer (2003a,b) and Biberauer & Roberts (2005), it becomes easy to state the GHC in a highly general form that no longer makes crucial use of the notion of linear precedence.

But before I move on to the restatement of the GHC, it is important to note a few basic characteristics about scrambling under the classical phrase structure. Under this view, scrambled elements regularly escape the c-command domain of heads that are potentially quite non-local, i.e. quite far up the structure from the base position (initial merge) position of the scrambling element. Thus, under the standard assumption that objects of verbs are initially merged as the complements of verbs[38] in VP or RootP, Japanese/Korean type scrambling potentially allows objects to escape the c-command of at least C, T, v, and V/Root, German scrambling moves objects above T, v, and V/Root, and so on. Additionally, each scrambling language allows some set of landing sites that are lower than the highest possible landing site (the position determined by the GHC, in the account I have given up to this point). Thus, a movement of this kind cannot be associated with any particular head's Probe for a feature on the scrambled object, either

[38] This generalization does not change under a system in which objects are initially merged at the specifiers of some lower verbal projection under little-v, as they would be in a system in the spirit of Marantz (1993). Scrambling still allows objects to climb quite a lot of structure.

cross-linguistically or within individual grammars. It follows that, modulo a system with grossly ad hoc feature placement for each possible string in these languages, scrambling is essentially unconstrained by the feature-checking mechanism. Given Minimalist assumptions about movement going back to the notion of "Greed" in Chomsky (1993, 1995), movement that is not triggered by the featural content of the moving element would be completely unconstrained if it existed, and so should not exist.

Clearly this is not a tenable approach to scrambling. However, scrambling appears to be a general operation, targeting multiple landing sites subject to semantic factors (e.g. scope, as in Lee & Santorini 1994, or definiteness as in Diesing 1997 *inter alia*), pragmatic factors (information structure), and prosodic factors (e.g. the scrambling of phonologically reduced pronouns, as in Cardinaletti & Starke 1999). I suggest that the generalization which I have termed the GHC will be missed in a syntactic architecture that cannot model this type of optionality directly, and would instead require us to obscure the nature of scrambling by ad hoc feature stipulations. Following Saito (1985), Webelhuth (1989), and Lee (1993; and cf. references therein), I understand scrambling to be a form of adjunction, and I would suggest that the optional nature of scrambling results from its status as a type of adjunction. In this way, scrambling is like modifier adjunction, differing only in the fact that the adjoined item is moved from an internal position and adjoined, rather than entering the derivation via adjunction. Also reminiscent of modifier-adjunction, scrambling targets a range of adjunction sites.

In order to allow for a proper restatement of the GHC, the phrase structure we adopt must have properties which are broadly in accordance with the modified version of

Kayne's (1994) antisymmetry hypothesis under bare phrase structure (as suggested in Chomsky 1995).[39] In particular, I follow Chomsky (1995: 420) in maintaining that (XP-) adjunction and movement to dedicated specifiers are different operations with different

---

[39] While this approach is compatible with bare phrase structure (in fact, it should remain intact regardless of whether all the details of BPS are accepted), one important clarification is in order concerning adverbials. The bare phrase structure approach in Chomsky (1995) treats syntactic words as terminals (which correspond to "morphemes" in the Distributed Morphology sense; cf. Halle & Marantz 1993, Embick & Noyer 2001, references therein, and much subsequent work by these authors), and since projection (i.e. "bar") levels are defined in relative terms, syntactic words can be simultaneously terminals and maximal projections if they happen to not combine with any other element during a derivation.

In the case of adverbials, this approach has the problematic consequence that adverbs which constitute their own adverbial phrases (i.e. one-word AdvPs) would be both heads and phrases, and therefore c-command out of the AdvP (c-commanding the vP, for instance), while the heads of multi-word AdvPs will not c-command out of the AdvP. Since the syntactic distribution of one-word vs. multi-word adverbials is generally the same, this is an unmotivated distinction in most cases. In the present context, this conclusion would allow the GHC to function normally in the case of multi-word adverbials to the left of vP, but cause, contrary to empirical evidence (e.g., from negation in Icelandic and Scandinavian generally), scrambling to be blocked just in the case that an AdvP left of the vP is made up of a single word. Since a full discussion of this detail of BPS is beyond the scope of this work, I will simply assume that AdvPs always contain more structure than just their head Adv, or at least, that they do in the canonical cases of adverbs that occur VP, vP, IP, or CP-adjoined and are bypassed by scrambling. I currently remain agnostic as to what constitutes this additional structure, and leave this as a question for further research, though I will mention two obvious possibilities. First, perhaps BPS can be weakened to allow unary branching structures in a few specific cases in which this characteristic is lexically specified (which is the same as saying that adverbs are inherently maximal). Second, it is possible that single-word AdvPs contain a null head, and the adverb is in fact the complement of that head. There is also the possibility that the latter two statements amount to the same thing: if AdvPs are always maximal, then they must always consist of two (syntactically-visible) morphemes, and so adverbs that project their own AdvPs are always internally complex from the point of view of the syntax (even though this may be obscured by a seemingly monomorphemic spell-out in PF). Whatever the solution is, the important point in terms of this discussion is that single-word AdvPs are not necessarily like clitic pronouns, which are truly minimal and maximal simultaneously, and must move from their base-positions to adjoin to some head (Chomsky 1995: 417-418). This must be case, or it would be impossible to have single-word adverbials with the same distribution as multi-word adverbials (e.g. vP-adjoined), because the single-word AdvPs would only ever surface in head-adjoined positions. The same problem arises for single-word pronominal DPs which are not clitics, "strong pronouns" in the sense of Cardinaletti & Starke (1999), and presumably these also have some internal structure that clitic pronouns do not have just as in the adverb case.

However, this is not to say that there don't exist true clitic adverbs which are truly minimal and maximal simultaneously and move to head-adjoin, just as there are true minimal/maximal pronominal clitics (Chomsky 1995: 405, 417). Old English *þa* (= "then") is a likely case (see Pintzuk 1996 for discussion of this adverb), and this analysis may be an explanation for the otherwise unusual distribution of a set of adverbials in English and Icelandic which appear following the subject and immediately preceding Tense. An example is English *never* in (i) below and Icelandic *mögulega* (= "possibly") in (ii) below:

(i) John wants to take weekends off, but he never can.

(ii) "…Synir og dætur hins mikla himnaföður," svara ég eins
sons and daughters the-GEN great heaven-father-GEN answering I as

types of landing sites, and that the distinction between specifiers and adjoined positions does not line up perfectly with the A/A' distinction (1995: footnote #18). Additionally, I accept Chomsky's conclusion that an antisymmetric approach is not incompatible with a system in which specifiers and adjoined positions are different. However, I differ with Chomsky in his brief suggestion that scrambling is a case of movement to a specifier, and instead follow Saito (1985), Webelhuth (1989), and Lee (1993) in taking scrambling to be a case of adjunction (see Lee 1993: section 1.3.2 and Lee's Chapter 5 for an extended argument to this effect). In fact, as will become clear in the discussion of binding in Chapter 8, scrambling is a core example of the A/A' distinction not matching up with the specifier/adjunction distinction: it is adjunction, but not A-movement (contra Lee 1993), nor A'-movement (contra Saito 1989), nor movement to a mixed position (contra Webelhuth 1989). Instead, the unusual properties of scrambling derive from its being the rare situation in which an XP is adjoined whose origin is internal rather than external (this is the normal case for head-movement). I will put off the full restatement of the GHC as a constraint on adjunction until I have made the antisymmetric account of head-finality explicit (see below), but informally, the constraint will state that adjunction cannot change the c-command relations between any given head and any other element during the course of the derivation.

Friendly á svipinn og ég mögulega get.
Friendly in expression-the as I possibly can.
" 'Sons and daughters of the Heavenly Father,' I answered, as much like Friendly as I possibly could."
(Hallgrímur Helgason 2008: 84; cf , also the discussion of this phenomenon in Thráinsson 1986)

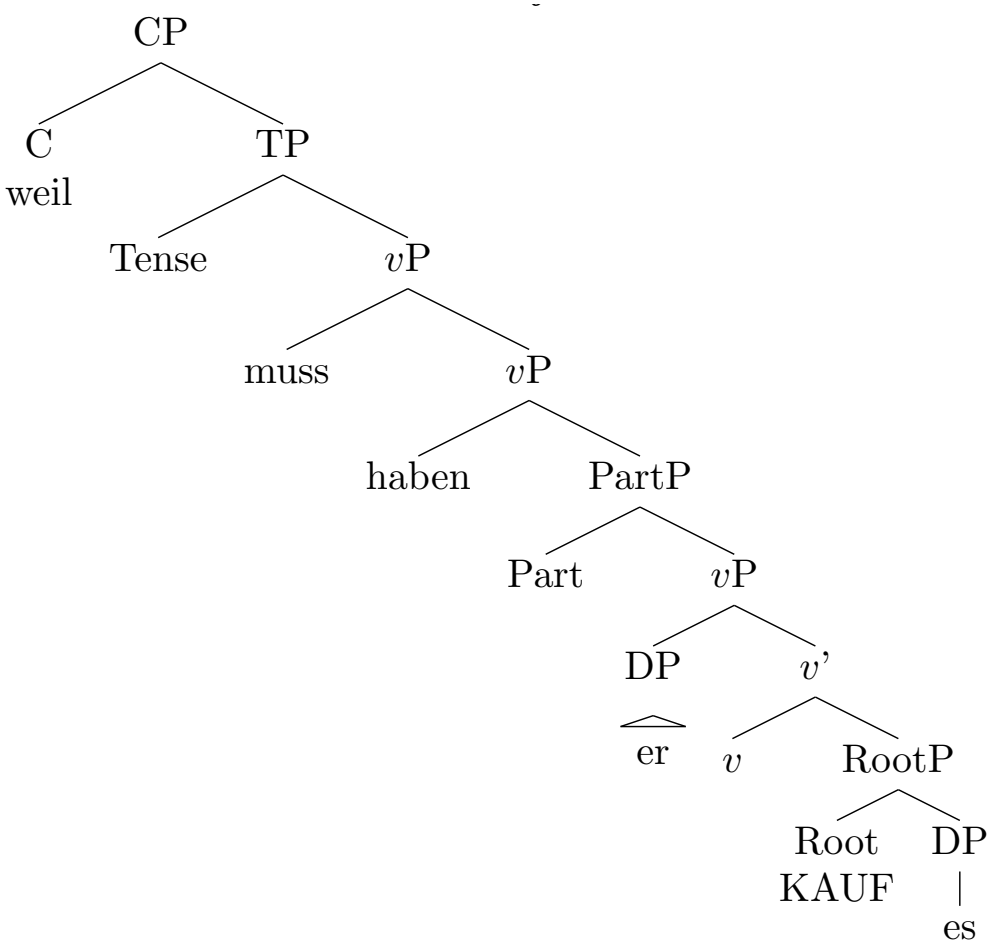
5.3 Pied-Piping and a Uniformly Left-headed Phrase Structure for German

Assuming the vP-hypothesis (Kratzer 1996) and that auxiliaries are functional verbal elements (i.e. "little-v"s, which head their own vP projections[40]), suppose the underlying structure of the (standard) German subordinate clause in (1) is as shown below in (2). (The German and Dutch examples throughout this chapter will be subordinate clauses, which allows us to examine how Tense-final order is derived without the complication of the V2 phenomenon, which is orthogonal to the issues at hand. V2 in German and Dutch matrix clasues has been standardly assumed to involve V-to-C movement to an uncontroversially head-initial CP since den Besten 1983). The remainder of this section will show that the initial-merge structure in (2) can derive the order in (1) with no ad hoc movements or feature assignments.

(1) …weil er es gekauft haben muß
because he it bought have must
"because he must have bought it"
(example from Wurmbrand 2004: 25)

(2)

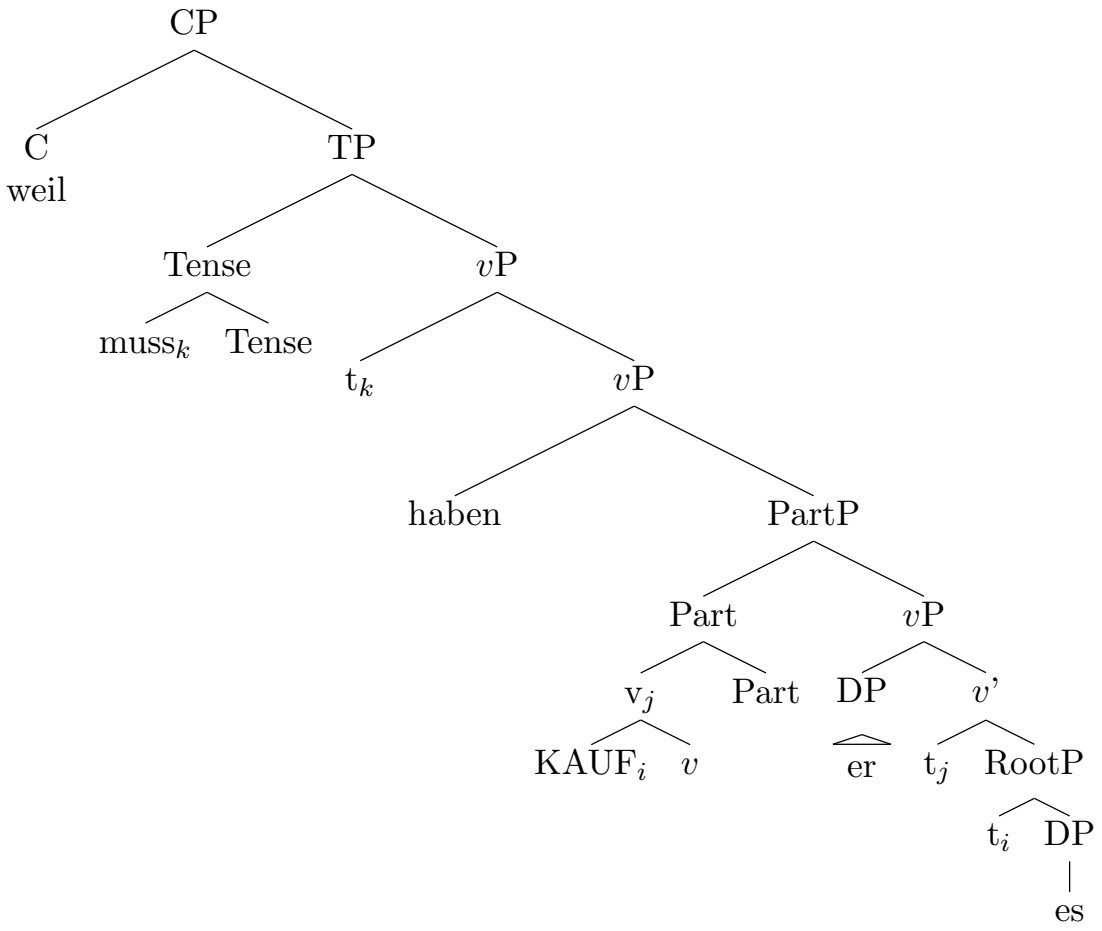
[40] It is possible that "vP" is not the best label for the projections headed by the various auxiliaries; they may in fact head Asp(ect)P(hrases) and/or MoodPs, depending on the auxiliary, but this detail does not affect the hierarchical properties of the structure and so is orthogonal to the discussion here. Similarly, I remain agnostic as to whether there may be more functional heads below Tense that the auxiliaries move to during the derivation. As long as the highest, finite verb ultimately lands in Tense and the auxiliaries land in positions below Tense but above the nonfinite lexical verb, assuming a more articulated structure should not have any effect on the proposals here.



In order to most clearly describe how the German surface order is derived, I will present each step of the derivation in turn, beginning with the head-movement component and then moving on to XP movement. The "RootP" in the structure above corresponds to "VP" in many notations: it consists of the internal argument and the yet unverbalized verb Root, which moves up to the next little-v to become categorized as a verb. It is this little-v which licenses the external argument of the verb, following (Kratzer 1996 and much subsequent work). The "PartP" is the functional projection of the features associated with participial morphology, and the now verbalized verb moves to incorporate with Part, with the result ultimately spelled out as *gekauft*. This sequence of

movements follows the assumptions of Distributed Morphology, in which the categorization of basic lexical items (or really, idiosyncratic lexical content without functional features) and additional morphology are generally the result of head-movement. Finally, the highest auxiliary moves to Tense, and the result of the head-movement stage is shown in the structure below.

(3)

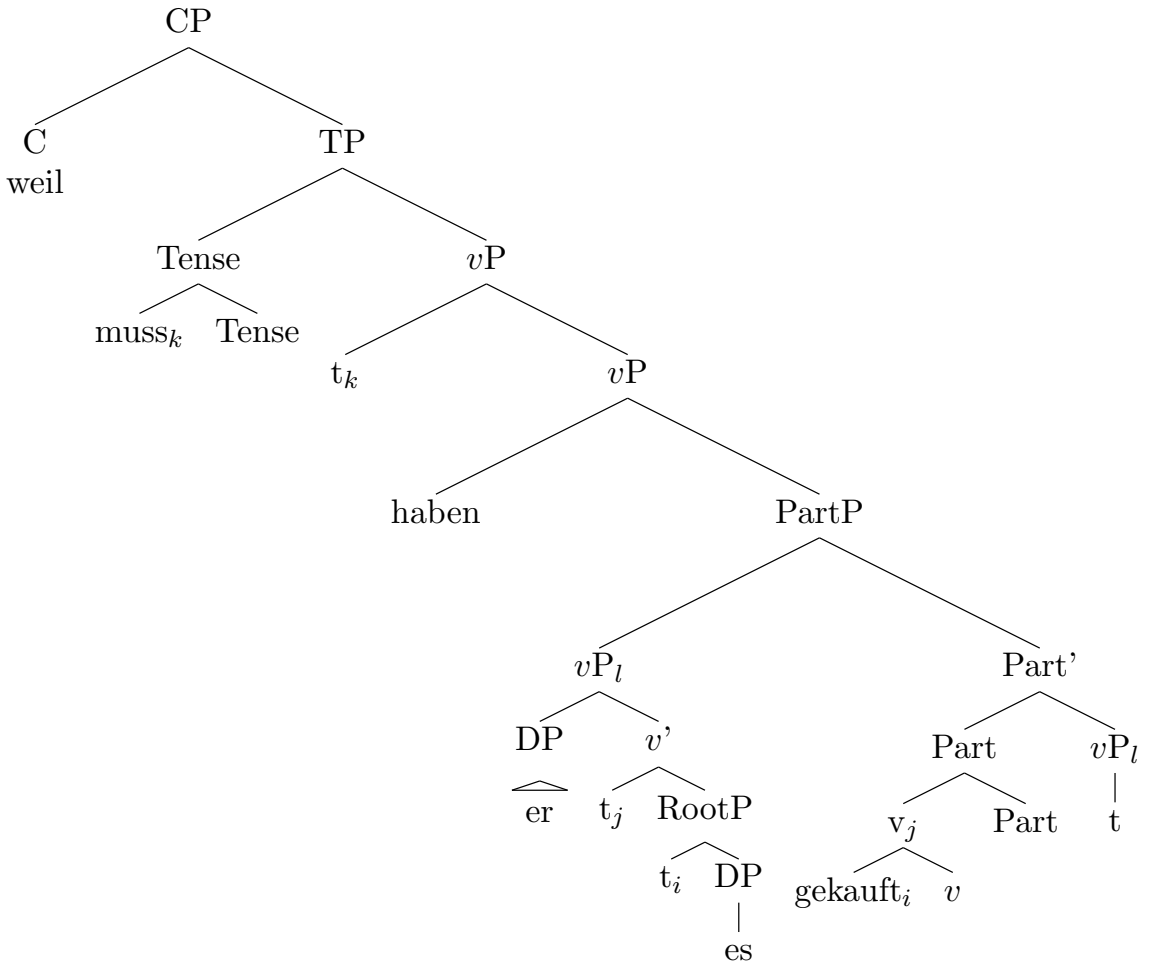


Continuing the derivation step by step, it is clear that some XP movement will be required in order to derive the German sentence from the initial Kaynian-style right-

branching structure in (2) above. The approach with the fewest unmotivated movements (and indeed, fewest total movements) is a "roll-up" derivation in which only maximal verbal projections move (i.e. vP, PartP); this follows the ideas in Biberauer (2003a,b), in contrast to "VP-evacuation"-style proposals (e.g. Hinterhölzl 2000, 2006, 2009). The latter type of approach requires vP movement to be preceded by the separate movements of most (or all) of the smaller constituents of the vP, each to a different specifier at the left of Tense. As I will discuss later in this chapter, the VP-evacuation type of approach is unattractive on conceptual (as well as empirical) grounds, requiring an unconstrained and largely ad hoc proliferation of specifiers in the T domain in order to derive all of the possible surface orders in a language like German (note that Kayne 1994: 141, in footnote #15, actually makes the same point concerning the idea that complementizer-final orders in, e.g. Japanese, could be derived by evacuation of the IP into the C domain).

The first step in the XP-movement component of the derivation is to move the lowest vP to Spec(PartP) after the nonfinite lexical verb has head-moved through the lowest little-v to Part; this stage of the derivation is shown in (4) below.

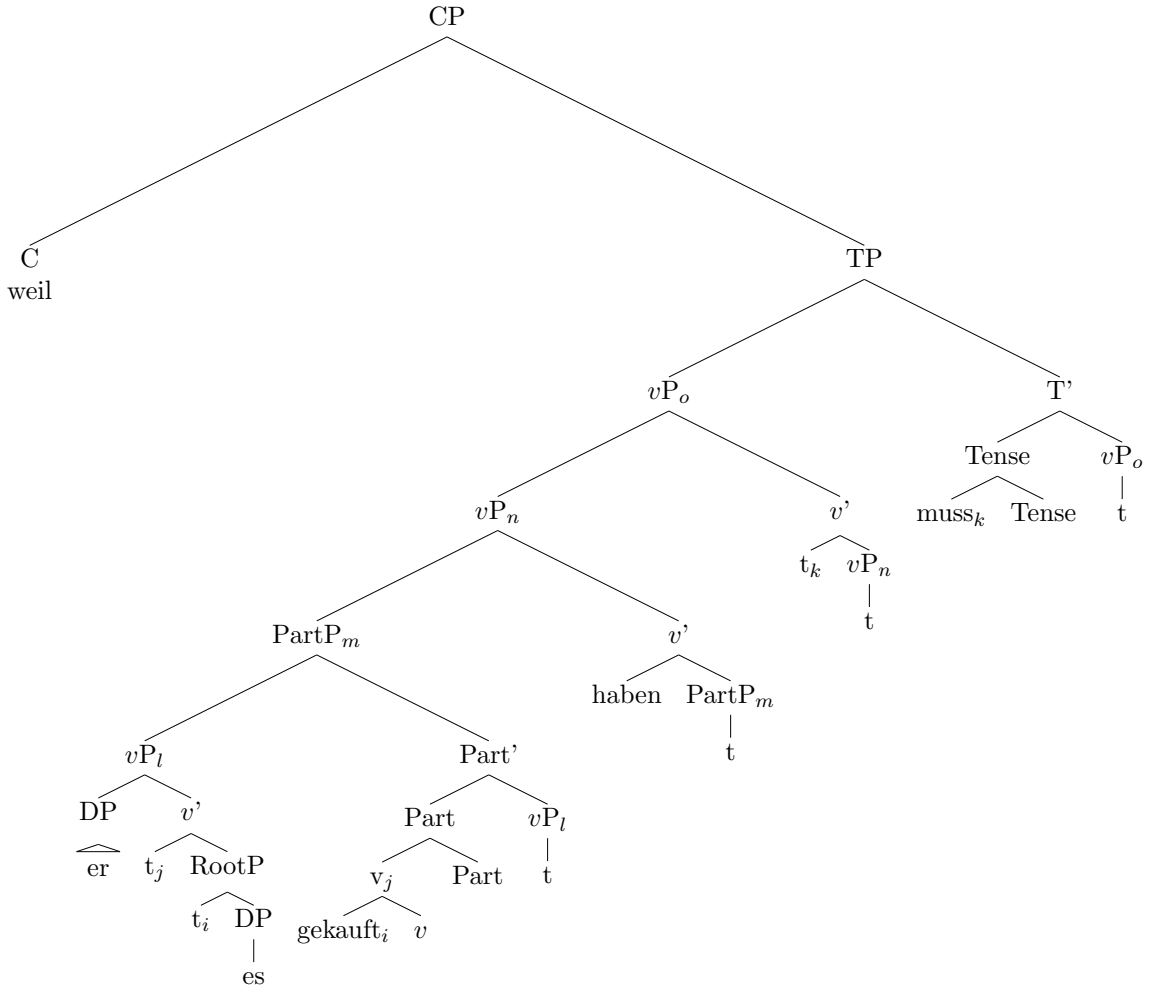
(4)



The last step in this procedure will be to move the highest vP to Spec(TP), as originally suggested (or in fact, implied) in Kayne (1994: 48) for Tense-final languages, but that operation follows a series of movements which result in the subject *er* being along the specifier line of TP (i.e. specifier of the specifier of the specifier of the specifier of the specifier of T) before the highest vP finally lands in Spec(TP). To derive the uniform head-final order in the German sentence in (1), after the lower vP moves to Spec(PartP), PartP moves to the specifier of the next highest vP, which in turns moves to the specifier

of the vP dominating it, and then finally the highest vP moves to Spec(TP). The complete derived tree is shown below in (5).

(5)



This is very similar to the proposal made for Old/Middle English in Biberauer & Roberts (2005), though it differs in a number of important ways (which will become apparent below). Most importantly, this analysis makes use of the important insight from Biberauer (2003b) and Biberauer & Roberts (2005) that the roll-up derivation is

ultimately motivated by the need to check an EPP feature on the Subject in Spec(TP). Note that in the derived tree above, the subject of the clause, *er*, lies directly on a line of specifiers from Spec(TP) to the specifier of the lowest vP (i.e. Spec(Spec(PartP))). This is not coincidental to the derivation, but following Biberauer (2003b), it is in fact the motivation for all of the movement that derives the correct ordering of verbs in German: the subject has moved to Spec(TP) in order to satisfy the EPP and to check nominative Case, and in doing so, it has pied-piped the maximum amount of structure with it. (Actually, to be more precise, the maximum amount of structure has been pied-piped by the **end** of the derivation. In fact, the first move up the tree by the subject only pied-pipes the vP projection of which it is a specifier, i.e., the projection immediately dominating it, as is usual for pied-piping. Then the second move up the tree pied-pipes one more projection, and so on until the structure in 5 is formed. In this way, pied-piping remains local to one projection up on each move). Thus, vP movement to Spec(TP) in a language like German checks the same EPP and Case features as the movement of subjects to Spec(TP) in English, though English differs from German in that the subject moves alone to Spec(TP). The difference between English, German, and, as I will show in the next section, Dutch, lies in the extent to which additional structure is pied-piped along with the subject to Spec(TP) and to each of the subject's intermediate landing sites on its way up the tree to Spec(TP).

However, the analysis in Biberauer & Roberts (2005) relies on the idea that vP fronts to Spec(TP) in **every** case that results in verb-finality on the surface. As they admit in a footnote #3, and this account as it stands can only derive the German

uniformly verb-final order in two sets of cases: first, where there is only one nonfinite verb (i.e. a single vP fronts to Spec(TP)), or second, when there are multiple nonfinite verbs and they plausibly head their own nonfinite clauses (i.e. each vP fronts to the specifier of its dominating TP). It does not extend to cases in which there are multiple nonfinite auxiliaries and they do not each project a full clause (TP).[41] Thus, in a sentence like the one in (1) above, Biberauer & Roberts (2005) are forced to claim that *muß* takes a clausal complement. The derivation shown above, on the other hand, takes Biberauer & Roberts' insight a step further and generalizes it to cases in which auxiliaries do not necessarily project clause-level structure. This step is particularly attractive in that it also allows the Biberauer-style proposal to be incorporated into the theory of Distributed Morphology. From a DM perspective, the participial morphology on the past participle *gekauft* in (1) must be associated with some functional head, and the combination of the morphology and the verb stem reflects head-movement of the verb (i.e. a verbalized Root) to the participial head.[42] However, it seems unlikely under standard assumptions that the functional projection associated with participial morphology is TP. In this way, the Biberauer-style analysis contains few ad hoc assumptions if the roll-up proceeds through Spec(PartP) in cases like (1), rather than an embedded Spec(TP).

---

[41] Or at least, where there is little or no evidence that each auxiliary projects a TP. In fact, the modification to Biberauer & Roberts (2005) I'm pursuing can be extended to the case in which a verb takes a clausal complement rather than a DP or vP, but it does not need to be. The exact structure that each type and form of auxiliary projects is an empirical question which I leave for further research.

[42] Thanks to Dave Embick for this suggestion. I will also briefly note that under DM assumptions, the participial morphology could combine with the verb as a result of PF lowering rather than head-movement. However, since German clearly has verb-movement in general, I see no reason to entertain this possibility at the present time.

Continuing this reasoning, we can show that the roll-up derivation does not necessarily require multi-clausal structure. If each auxiliary is taken to be a little-v (i.e. a purely functional verbal element), they each potentially project their own specifier. Taking the little-v hypothesis of Kratzer (1996) to one of its logical consequences, every specifier of a verbal element is a potential position for licensing an external argument, whether it is externally-merged (as in the usual case with the lowest vP), or internally-merged (as with raising verbs or unaccusatives).[43] If this is accurate, then the Minimal Link Condition[44] forces the subject to move cyclically up the tree: the subject (or the subject-containing vP, PartP) cannot skip any potential landing sites as it is attracted by Tense up to Spec(TP), so each movement of the constituent containing the subject must land in the next available specifier position (Spec(vP) position), and in (standard) German sentences like (1), the largest possible constituent is pied-piped with each move up the tree. In this way, the only EPP feature driving the derivation is the one on Tense, which is a familiar characteristic from well-known languages like modern English. This

---

[43] I do not, of course, mean to suggest that every Spec(vP) could be filled by a different externally-merged subject, deriving umgrammatical sentences like:

(iii) * John must Mary have Bill bought it.

Such sentences would be an obvious violation of the theta-criterion. In the usual case, the subject of the theta-assigning verb will move cyclically through each Spec(vP) when it is attracted to Spec(TP), in a similar fashion to raising-verb constructions.

[44] I use the term "Minimal Link Condition" to make it clear how the movement is constrained; I do not mean to argue against the idea that MLC can be incorporated into a statement of "Attract", as proposed in Chomsky (1995). But for the purposes of this proposal, the key point is that the subject (with or without pied-piped vP structure) cannot skip potential landing sites in Spec(vP) or Spec(PartP) as it is being attracted up the tree by Tense, just as wh-movement must move through intermediate Spec(CP)s in a multi-clausal structure when the matrix C is the attractor for the wh-word. I will also note that the latter phenomenon has a natural explanation under phase theory, and so it is likely that the subject movement (and vP, PartP) movement I describe here could also be accounted for under phase theory. However, it is not crucial to the discussion here whether or not my analysis should be recast in those terms, or if it were, whether it would be necessary to adopt all of the details of the Chomsky 's (1998/2000) phase theory.

feature at the top of the tree must be satisfied by movement of the subject, and that movement must proceed through any potential intermediate landing sites on the way to Spec(TP). The chain of intermediate landing sites is nothing unusual or specific to this problem: it is created by the same mechanism as cyclic wh-movement. In languages like standard German, there is then the additional requirement that each movement of the subject pied-pipe one additional level of structure (exactly how this is formally encoded will be fleshed out below).

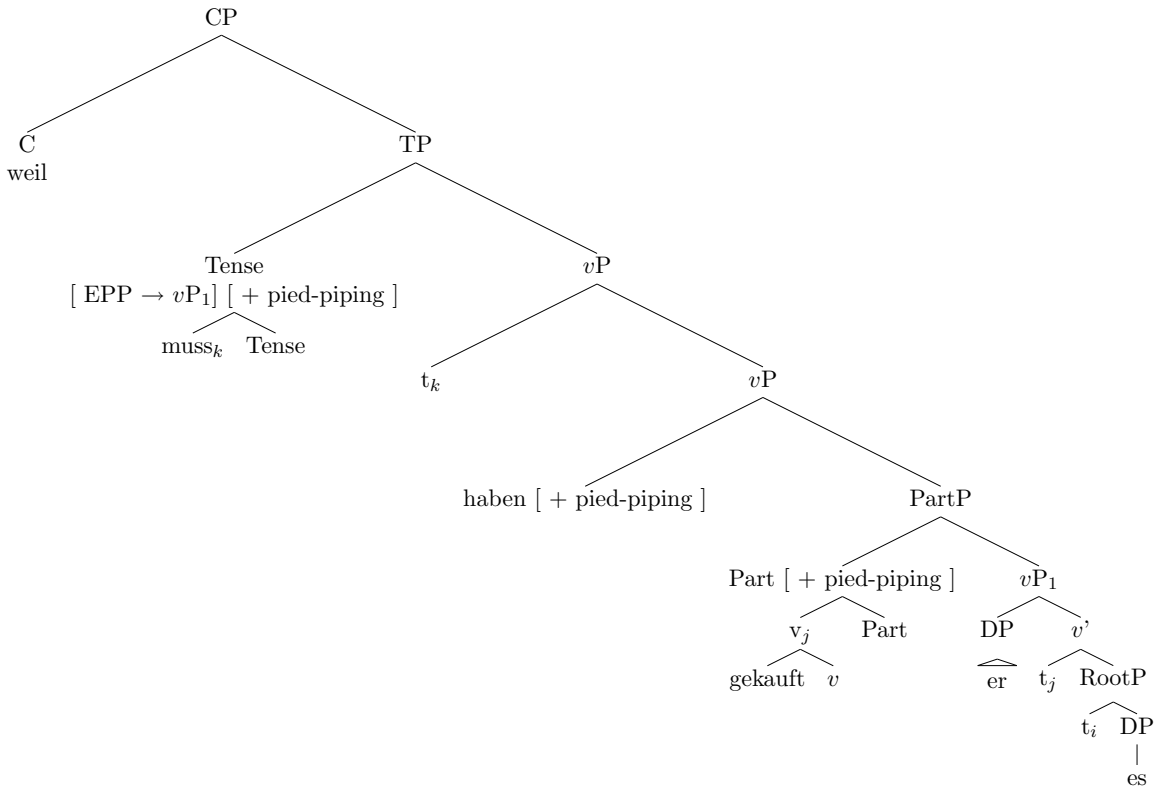
At this point, it is important to make more explicit the formal mechanism which controls the pied-piping of subject-containing constituents, and once again, I being with a discussion of the intellectual predecessor to the current analysis in Biberauer & Roberts (2005). In order to account for the cross-linguistic variation between "Tense-final" and "Tense-initial" languages, Biberauer & Roberts (2005, esp. Table 1, Table 2) propose that each language has a [± pied-piping] feature on Tense, and the difference between languages like German and English is due to the former having a [+ pied-piping] feature-value on Tense and the latter having a [- pied-piping] feature-value. The parameter separating Tense-final and Tense-initial languages is that the former combine their EPP feature with pied-piping, with the results that they attract the vP with the subject in its specifier to Spec(TP), while the latter attract only the subject DP itself to Spec(TP). But again, this important idea can be extended to cover a larger range of data if we carefully consider what it means for [± pied-piping] to be a feature of functional heads, and what consequences that might have for understanding cross-linguistic variation in [± pied-piping].

First, on purely conceptual grounds, it cannot be the case that languages, per se, are parameterized for [± pied-piping]; rather, it is individual functional heads which carry this feature and are specified for some particular value. This conclusion is inevitable in the light of research on "competing grammars" within individuals, which shows that bilingualism (or bidialectalism), of a sort, can exist even when the "languages" in question only consist of different values for some syntactic parameter (Kroch 1989, Pintzuk 1991, Santorini 1992, 1993a). The idea that individuals can exhibit more than one setting for a given syntactic parameter has led to a precise understanding of the syntactic variation that accompanies syntactic change, and an understanding of its inherently unstable nature. In situations of phrase structure change, individual speakers always exhibit both structural variants in their linguistic performance during the time period in which the change is in progress, even though the given state of variation ultimately turns out to be unstable over time (this has been shown in the studies cited above, and especially highlighted in Santorini 1992; note that this is also true of every sound change in progress that has been studied to date; cf. Labov 1994 and references therein). Kroch (1994) explains the contrast between synchronic intraspeaker variation and diachronic instability by proposing that the locus of syntactic variation is in the featural content of functional heads, which allows the problem of syntactic variation to be reduced to the phenomenon of morphological doublets (e.g. *dive ~ doved*). In both cases, individual speakers are able to use functionally equivalent, formally distinct elements, but there is also a cognitive pressure against these doublets ("blocking") that eventually prevents the doublets from being transmitted stably over generations.

If this view is correct, the locus of syntactic variation is not the "language" or even the idiolect, but the functional head. Once we make this step, the [± pied-piping] parameter must be a feature of functional heads, and there are three immediate consequences for the antisymmetric theory of verb-final languages. First, there is no *a priori* reason to think that the [± pied-piping] feature applies only to the Tense head; it could very well be a feature that must be specified for all heads along the clausal spine. The standard German sentence in (1) would then have the following (relatively uninteresting) pied-piping-feature specifications for verb-related clausal heads; each head is [+ pied-piping], causing one additional level of structure to be pied-piped with each XP movement up the tree. (The structure below is shown after head-movement, but before XP movement; a version of (3) above.)[45]

[45] The specification in the EPP feature is my adaptation of the original claim in Biberauer (2003b), Biberauer & Roberts (2005): that the initial attract targets the vP containing the subject at initial merge, vP_1. Again, this specification of pied-piping in the EPP feature is the parameter that differentiates Tense-final from Tense-initial languages, whether or not all of the other [pied-piping] features in the clause are "+" or "-" (as we will see, German and Dutch will differ on this point, though they are both Tense-final in the EPP-pied-piping sense).

(6)



Note that, in a sense, Tense-final languages can potentially have two types of [pied-piping] features associated with the Tense head. One is really part of the EPP feature: it determines how large the minimum constituent must be to satisfy the EPP by occupying Spec(TP). Following Biberauer & Roberts, all Tense-final languages are specified for a [+ pied-piping] version of the EPP, attracting a vP-sized constituent containing the subject. However, in addition, each functional head is specified for whether or not it will allow movement to its specifier without the pied-piping of its complement. (The formal definition and empirical consequences of this general pied-piping feature will be explained below.)

The second consequence of assuming pied-piping as a general feature of functional heads is actually just a corollary of the first: if the feature is specified for all heads, then it is possible that some verbal heads in the clause could be specified for [+ pied-piping] while others are specified for [- pied-piping]. One example of this would be "SIOV" languages, like the Vata and Gbadi (Koopman 1984) and, as I have argued in Chapters 2 and 3, Yiddish. These languages have [- pied-piping] on Tense (both in terms of the EPP and in terms of the general pied-piping feature on functional heads), but [+ pied-piping] on other verb-related heads. Again, the need to attract the subject to Spec(TP) to satisfy the EPP feature on Tense and check nominative case is the same as in German or English. The difference lies in a slightly different distribution of pied-piping-features among the functional heads along the clausal spine. If we understand the [± pied-piping] distinction to apply to all of the functional heads along the spine of the clause, just like any other feature associated with a given functional head, then it is entirely expected that a given language variety might show a number of heads with different specifications for [pied-piping].

Before moving on to further empirical consequences of the antisymmetric, pied-piping analysis, I would like to clarify some formal characteristics of the pied-piping mechanism. I mentioned above that pied-piping operation is intended to be local; in order to prevent massive overgeneration of various kinds, each movement of an element attracted by the EPP should only be able to pied-pipe one additional level of structure at a time. While this assumption about pied-piping is not usually discussed explicitly, it is familiar from the canonical case of pied-piping in wh-movement. If an arbitrary amount

of structure could be pied-piped, then we might expect a wh-word within a PP to pied-pipe an entire TP, deriving the ridiculously ungrammatical (7c) from (7a) below, rather than the grammatical (7b):

(7) a. John said that Mary gave a book to who?
b. To who did John say Mary gave a book?
c. * Mary gave a book to who did John say that?

The requirement of locality in pied-piping is also at work in examples like those in the English and German contrasts in (8) and (9) below (cf. Bayer 2006, discussing in part earlier observations by Webelhuth 1989 and van Riemsdijk 1985).

(8) a. I wonder [$_{+wh}$ [$_{+wh}$ whose books] we should read]
b. * I wonder [$_{-wh}$ [$_{-wh}$ books by whom] we should read]

(9) a. Ich frage mich [$_{+wh}$ [$_{+wh}$ mit wem]$_1$ [es t_1 aufzunehmen]]$_2$ dir t_2
I wonder REFL with who it up-to-take for-you

ein Bedürfnis wäre]
a need would-be
"I wonder with who to enter into a competition would be a need for you"

b. * Ich frage mich [$_{-wh}$ [$_{-wh}$ [es mit wem aufzunehmen]]$_2$ dir t_2
I wonder REFL it with who up-to-take for-you

ein Bedürfnis wäre]
a need would-be

c. * Ich frage mich [$_{-wh}$ [dir [es mit wem aufzunehmen]
I wonder REFL for-you it with who up-to-take

ein Bedürfnis wäre]
a need would-be

(Both 8 and 9 adapted from Bayer 2006: example 21)

In order to pied-pipe a larger constituent containing a wh-expression, as in the English DP in (8) or the German infinitival clauses in (9), the wh-expression must generally occur at the left edge of the constituent to be pied-piped.[46] This observation is easily accounted for if pied-piping can only see "one level down", and the wh-phrase in the grammatical examples above occupies, or has moved into, the specifier of the larger constituent. The locality constraint on pied-piping can be formalized in the following way:

(10) Definition: **Pied-Piping**

Suppose a functional head, H, has a complement, XP, and some phrase, YP, is in Spec(XP).[47]

---

[46] However, as Bayer (2006) notes in footnote #4, there is not a clear consensus as to why the following example from Ross (1967) is grammatical:

(iv) the report [[the height of the lettering on the covers of which]$_i$ the government prescribes t$_i$]

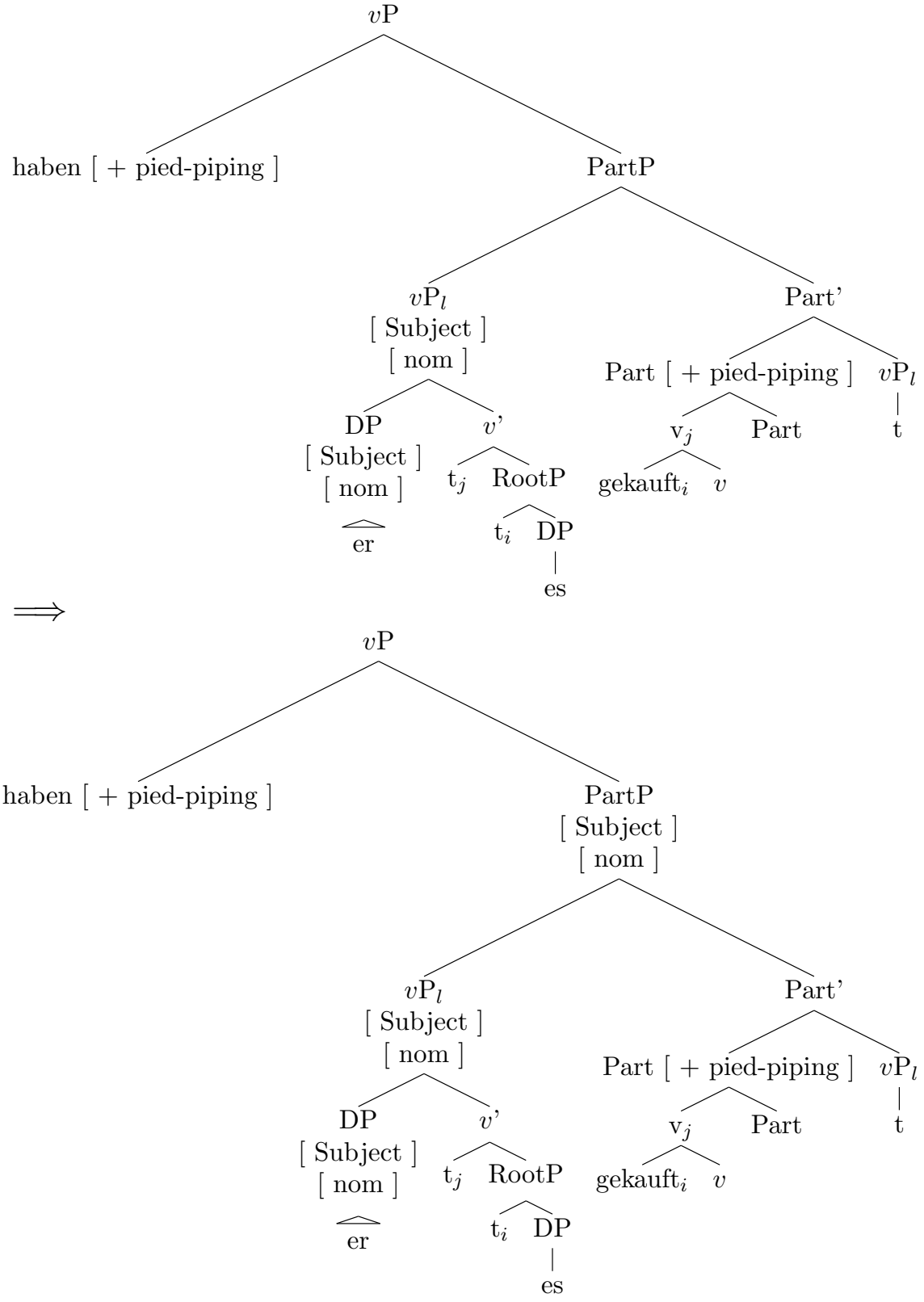
A full discussion of why pied-piping seems to be less restricted in some cases of English relative clauses would be beyond the scope of this dissertation. The important point for the discussion at hand is that pied-piping must be more restricted in most cases than it is in the sentence from Ross, and I will assume it is more restricted in the Tense-final clauses under discussion here.

[47] All of the cases I will discuss involve pied-piping a constituent because something in its specifier has been attracted. I did not include the "Comp(XP)" case in the definition above so as not to complicate the discussion, but ideally, the definition should be stated to also include phenomena such as PP pied-piping with wh-movement. However, this type of pied-piping seems to be more restricted than pied-piping by the specifier in some languages (e.g. English, as I will mention briefly in a few paragraphs in my discussion of McCloskey 2000), and I do not currently have any formal proposal to make concerning the difference between the two. It may be that the latter case, pied-piping by a complement, is really only possible when there is no specifier present and there is no mechanism to create the specifier, as will be the case for most English PPs under a bare phrase structure. Note that some English PPs, such as *thereof* , and its counterpart in German and Dutch, such as *davon* and *daarvan* ("thereof"), respectively, fall into the usual pattern described here, in which the attracted element must move to the specifier of its containing maximal projection before pied-piping can take place.

H[+ pied piping] implies: *value*(F, XP) = *value*(F, YP) when:

1. YP is attracted by a Probe higher in the structure than H, probing for feature F with *value* α
2. *value*(F, YP) = α (i.e. YP has the feature specification [α F]
3. XP is unspecified for [F]

For instance, the *haben* little-v head continues the feature percolation to PartP in the German derivation in the way shown below:


vP
haben [+ pied-piping]
PartP
vP$_l$
[Subject]
[nom]
Part'
DP
[Subject]
[nom]
er
v'
t$_j$
RootP
t$_i$
DP
es
Part [+ pied-piping]
vP$_l$
v$_j$
Part
t
gekauft$_i$
v
⟹
vP
haben [+ pied-piping]
PartP
[Subject]
[nom]
vP$_l$
[Subject]
[nom]
Part'
DP
[Subject]
[nom]
er
v'
t$_j$
RootP
t$_i$
DP
es
Part [+ pied-piping]
vP$_l$
v$_j$
Part
t
gekauft$_i$
v

Once the feature-copying in (10) has applied, the actual pied-piping is simply a consequence of Closest/Shortest[48] Attract: the maximal projection containing the original target of the Probe, which now has the appropriate feature specification, will move rather than the original target.

Conversely, a head specified for [- pied-piping] will have the opposite effect from what is described in (10). A [- pied-piping] head, finding a certain feature value [α F] on the specifier, XP, of its complement, will:

1. Check Spec(XP), YP, for the same feature-value pair [α F].
2. If it finds *value*(F, XP) = α = *value*(F, YP), then delete F on XP.

And of course, if a given head is not specified for [± pied-piping] at all, then it is simply inert; it will have no effect with respect to pied-piping in the case that its complement moves.

To complete the discussion of the formal nature of pied-piping, it is necessary to clarify a statement I made at the beginning of the pied-piping discussion, namely that Tense behaves as if it can pied-pipe 2 levels (in contrast to the normal case of pied-piping 1 level described in the definition in 10). Tense in German pied-pipes its vP complement in the roll-up derivation, exactly as (10) prescribes, but its EPP feature is also looking to attract a vP rather than a DP(-subject) to begin with. Thus, Tense appears to exceptionally have 2 [± pied-piping] features, one behaving as it does on all verbal heads, and a second one associated with the EPP. As I mentioned above, this characteristic of

[48] "Closest" here must be computed in terms of number of nodes traversed on the path from the Probe to potential Targets.

Tense is the parameter difference between Tense-final languages and Tense-initial (i.e., Tense-medial) languages, and it currently appears like a glaring exception to the way the rest of the pied-piping mechanism works. That is not strictly speaking true, however: Tense is only inherently associated with the [± pied-piping] feature that it has by virtue of its being a Probe, the one on the EPP. Because Tense is a Probe, its [± pied-piping] feature is different from the pied-piping feature on other heads in that it is a feature of a feature; it is a feature of the feature that makes Tense a Probe: the EPP. The fact that Tense can additionally cause the pied-piping of its complement vP is a consequence of head-movement. Normally, some lower verbal head must move to incorporate with Tense, as *muß* does in the German structure in (2). When this occurs, all of the features of *muß* incorporate with the features on Tense. If *muß* is specified for [+ pied-piping], then this feature transfers to the Tense head, and Tense will pied-pipe its complement like any other head, in addition to its inherent ability to pied-pipe vis-à-vis the EPP.

The inherent pied-piping of Tense manifests itself as a feature of the EPP feature (represented in 6 above as "EPP → *v*P"), in accordance with the present system in the following way: the EPP feature on Tense probes down the structure for its target, and it looks down the tree from head to complement (down each sister). When it arrives at the last head in the tree whose complement contains the sought-after element, an EPP with [+ pied-piping] copies the features of the element on to its maximal projection as in (10). The special status of "Probe" in the system allows the pied-piping feature from the EPP to incorporate with the features of the last relevant verbal head on the clausal spine.

Admittedly, this is an added complication to the system, but it is no more complicated than the combination of Attract and (10), and it is a necessary distinction between Tense-final and Tense-initial languages which is not any more stipulative than the statement that "TP is right/left-headed". This added layer of pied-piping may also not be necessary to derive the standard German sentence in (1), but as will become clear in the next section, it is necessary to derive Tense-final languages with the West Germanic verb-(projection)-raising construction, such as Dutch and Swiss German. Additionally, I suggest the following theorem:

(11) **Theorem:** Feature copying (e.g. by head-movement/head-incorporation) cannot lead to feature bundles with conflicting feature-value pairs.

In addition to being necessary under any reasonable theory of features, this theorem makes intuitive sense in this context: no head can cause features to copy and percolate in its complement while at the same time deleting the same percolated features.

Finally, before moving on to the empirical consequences of the formal approach to head-finality, it is important to note that the type of pied-piping which derives the German head-final order is independently attested in the realm of wh-movement. Assuming that the understanding of the examples in (8) and (9) I've outlined above is correct, then those sentences constitute evidence that pied-piping can be triggered by elements in the specifier of a projection, not only in the complement (e.g., as in the pied-piping of PPs by wh-elements in German and English). Already we can see that the pied-piping necessary to derive the head-final orders of standard German is of a well-known

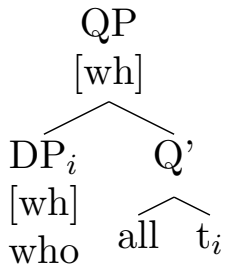
type. An even clearer case of pied-piping by a specifier is found in American and Ulster English sentences such as (12) below.

(12) [Who all]$_i$ did you meet t_i when you were in Derry?

(13) [What all]$_i$ did you get t_i for Christmas?
(McCloskey 2000: 58)

These examples, as analyzed in McCloskey (2000), also involve movement of the attracted element into the specifier of the maximal projection that contains it (see also the references in McCloskey 2000: 59 to this approach to other types of English quantifier float, beginning with Sportiche 1988). In fact, the movement of the wh-word into the specifier of the dominating QP is necessary in order to license pied-piping of the whole projection (2000: 60).[12] McCloskey gives the structure below for the wh-QP, once the wh-word has moved to Spec(QP) and the [wh] feature has percolated to the QP level (I have replaced his "DP" label for the maximal projection with "QP" below.) Note that this is the same feature-copying mechanism I described in (10), though McCloskey does not give a formal account of how the feature-copying proceeds.

(14)



In standard English (or, at least in the variety of American English which accepts examples like 13 and 14), the copying of the [wh] feature has the result that Closest/Shortest Attract causes the entire QP to move as far up the tree as possible, to the highest Spec(CP). Thus, (15) below is ungrammatical for many English speakers, where the lower [wh] marked element, the DP *who*, has moved rather than the whole QP.

(15) * Who did you meet all when you were in Derry?

McCloskey does not give an account of why (15) is ungrammatical for the relevant dialect(s) of English, mostly because his primary goal is to account for why it is grammatical in mid-Ulster English, along with more complicated examples like those in (16):

(16) a. **What**$_j$ do you think [t_j **all**]$_i$ (that) he'll say t_i (that) we should buy t_i ?
b. **What** do you think (that) he'll say **all** (that) we should buy?
c. **What** do you think (that) he'll say (that) we should buy **all**?
(McCloskey 2000: 62)

McCloskey's analysis of the options displayed in (16) relies on the idea that Closest Attract does not calculate "closest" by calculating the distance between the Probe and target in terms of number of nodes, but rather in terms the (asymmetric) c-command relations between potential targets (2000: 60). Unfortunately, although this idea makes the correct prediction for mid-Ulster English, it does not explain how a grammar of English can reject (15) or (16), which clearly are ungrammatical for some speakers. It predicts the mid-Ulster case, in which the lower wh-phrase can always be moved at any

point in the derivation, stranding the rest of the QP at any intermediate-trace site, but it wrongly predicts that standard/American English cannot exist: i.e., a variety in which the whole QP **must** be pied-piped along with the wh-phrase. This unpleasant consequence comes from locating the mid-Ulster pattern in the definition of Closest Attract itself, rather in the feature-copying mechanism which determines the possibilities for pied-piping.

If, alternatively, we adopt the proposal put forward in this chapter that pied-piping results from functional heads copying or removing features, then it is possible to account for the variation in (16) and the non-Ulster English judgment in (15) without having Closest Attract work differently in the two varieties. Suppose Closest/Shortest Attract calculates "closeness" based on the number of nodes traversed from Probe to potential targets, as I have suggested above. Then, in a structure like (14), Closest/Shortest Attract will always cause the whole QP to move, and *all* will not be stranded in intermediate positions. This is the standard/American English case, in which (15) and (16) are bad, and the pattern in (12) and (13) is the only possible option. This results if matrix (Question) Comp in standard/American English has a [+ pied-piping] feature, which, because Question-C is a Probe, will really be a feature of the [wh] feature on Comp. The status of matrix C as a Probe causes the [+ pied-piping] feature to travel down the tree and copy the [wh] feature to just one level higher than the wh-expression, producing the structure in (14). This is exactly parallel to the way the [+ pied-piping] feature associated with the EPP on Tense operates in Tense-final languages. If there are no other heads in the clause that are specified for [± pied-piping] and are also along the path of the QP as it

moves cyclically up the tree, then the only possible result is the standard/American English pattern in (12) and (13): exactly one level of structure is pied-piped. As we will see in the next section, this is also the way in which verb-projection-raising languages are Tense-final: they will end up pied-piping exactly one vP, as I will show below.

For the mid-Ulster case, the various possible patterns can be captured if matrix Question-C has a [- pied-piping] feature, in contrast to standard/American English, and so it will delete a copy of the [wh] feature at some point in the derivation, if possible. Then, the default case, if no other heads along the wh-expression's path to matrix Spec(CP) are specified for [± pied-piping], is the pattern in (16c): *all* is stranded in the base position. However, if there is variation in mid-Ulster English in the pied-piping abilities of lower heads in the clause, e.g. *buy* in (16), other options become possible.[49] If *buy* is specified for [+ pied-piping] but matrix C is [- pied-piping], then the QP will be pied-piped once up the tree, but then stranded as the wh-expression moves alone to matrix Spec(CP). This is the pattern in (16b). In this way, the pied-piping patterns described in McCloskey (2000) are plausibly derived by the same mechanisms that derive OV and Tense-final languages.

5.4 The Conservation of C-Command (Scrambling and the GHC, Revisited)

[49] I currently have no evidence bearing on the question of whether the variation in functional heads which can be either [+ pied-piping] or [- pied-piping] in mid-Ulster English is a case of competition between morphosyntactic doublets, and hence unstable diachronically, or whether the different versions of each head are distinct enough in usage (based on semantic, pragmatic, or even prosodic factors) that they are not in competition. In principle, either state of affairs could hold in this case.

The adoption of Kayne's (1994) conception of head-finality, in which all head-final structures are derived by the leftward movement of the given head's complement, allows the GHC to be rewritten entirely in terms of hierarchical relations without direct reference to linear precedence. Once the antisymmetric approach to head-final languages has been spelled out, it is possible to see that scrambling (with "object shift" as a subcase), is only sensitive to c-command, and the difference between heads appearing to the left or right of potentially scrambled elements is entirely derivable from the c-command relations. The GHC can also be seen as a general constraint on the operation of adjunction in general, not a specific constraint on scrambling. I have renamed the constraint below, in accordance with its more general scope of application:

(17) **Conservation of C-Command:**

Adjunction cannot subtract a c-command relation holding between a head and a non-head.[50]

(E.g., Scrambling cannot change the hierarchical relationship between head and phrase, but other operations, like A and A'-movement, can do so.)

This constraint, coupled with the following hypothesis about the nature of scrambling, derives the GHC:

[50] By "head", I specifically mean a "MWd", defined in Embick & Noyer (2001: 574) as a head not dominated by further head-projection. This excludes sub-parts of complex heads and traces of head-movement (which results in head incorporation).

1) Scrambling is syntactically optional (though it may be mandated by semantic, prosodic, or information structural concerns).
2) All optional operations (i.e. scrambling, modification) are instances of adjunction.[51]
3) Adjunction can be via external/initial merger (e.g. modifier-adjunction) or internal merger.
4) Scrambling is internal adjunction.
5) Therefore, scrambling obeys the Conservation of C-Command like other types of adjunction.

I mentioned about that scrambling is an unusual case of internal adjunction, where an element is adjoined after being initially merged in another position (e.g. in object position).[52] Aside from scrambling, adjoined phrases (e.g. PP modifiers or AdvPs)

[51] Note that I am not claiming the reverse, that all adjunction reflects an optional process, though it may very well be the case that all non-head adjunction reflects optional processes.

[52] However, internal adjunction is the normal case for head-movement (following Roberts 2001, references therein, and especially the incorporation analysis of head-movement in Baker 1988). For this reason, it is extremely tempting to replace "non-head" with "element" in the definition in (17), and assert that the Conservation of C-Command can hold between two heads, as well as between heads and phtases. On the face of it, this would be a welcome result, since it would then be possible to derive the Head Movement Constraint (originally stated in Travis 1984) from the Conservation of C-Command, making illicit head-movement and illicit scrambling subcases of the same restriction. Under the assumption that head-movement is a type of adjunction, when a head moves, it must adjoin to another head, but it cannot leave the c-command domain of the next highest head in the tree. Therefore, its only available landing site is the next highest head itself: by (17), head-movement cannot skip c-commanding heads.

However, stating (17) in a way so as to include the HMC unfortunately appears to be too restrictive. First, syntactic cliticization has been widely considered a type of head-movement since at least Kayne (1991) and Chomsky (1995), but it does not necessarily obey the Head Movement Constraint: as Kayne (1991) argues for a number of Romance dialects, clitic pronouns can frequently head-move to pass V/v and adjoin to Tense in a number of contexts; something about their featural content, the featural content of Tense, and the (apparently irrelevant) featural content of V/v makes the necessary movement

actually enter the derivation by adjunction. This is probably the reason that scrambling has never fallen neatly into the A/A' typology of movement (as I discuss in Chapter 8); it is an unusual type of movement, as it results in adjunction, and it is also an unusual type of adjunction, being derived by movement.

Note also that the statement in (17) derives the original "Holmberg's Generalization" of Holmberg (1986) for the Scandinavian "object shift" languages: objects can scramble as far as the head that immediately c-commands them. If that head incorporates into another head by head-movement (following Baker 1988 and much subsequent work), then moving the object farther than the head's original position does not remove the c-command relation between the head and the scrambled element as long is it is not scrambled past the new complex head.[53] For the purposes of (17), I assume that traces of head-movement do not increase the total number of c-command relations in the tree; in other words, the trace does not constitute a second c-command relation between the same head and the nodes it c-commands. However, it is possible to to avoid the problem of traces altogether by restating the Conservation of C-Command in terms of the copy theory of movement (Chomsky 1993), in the following way:

---

directly to Tense without stopping at V/v. This may be predicted under an appropriate interpretation of Relativized Minimality (or its reformulation as Closest/Shortest Attract), in which the minimal domain for attraction of an element is determined based on the featural complexes of the intervening elements (as in Roberts 2001, Rizzi 1990, Rizzi 2001). Indeed, if the HMC can in fact be reduced to Relativized Minimality, as Rizzi asserts, then there is no need to try and reduce the HMC to the Conservation of C-Command, since Relativized Minimality is independently necessary as a component of the grammar. Additionally, there is evidence from Breton that there can be "long head movement" of verbs in some cases (Borsley et al 1996; also see the summary of work on long head movement in languages other than Breton). Under certain conditions, verbs in Breton are able to move and skip intervening heads, provided that the intervening head does not have the appropriate featural specification to be a landing site for the verb (cf. Roberts 1994). This fact can be accommodated under a precise, feature-based formulation of Relativized Minimality or Closest Attract, but would be ruled out if the Conservation of C-Command applied to head-movement under its current formulation.

[53] I assume, I think uncontroversially, that the resulting complex head ("MWd" in Embick & Noyer 2001: 574) inherits the characteristics of all of the incorporated H^0 heads.

(17') A phrasal adjunct and all of copies resulting from the adjunction must be c-commanded by the same functional head.

According to the version of the constraint in (17'), traces of head-movement can clearly have no effect on scrambling. As long as the same functional head, including a complex head of which it is a part, c-commands all of the copies of a scrambled phrase at the end of the derivation, the fact that a trace of the head is lower than some of the copies is not relevant.

The statement of the Conservation of C-Command in terms of the copy theory of movement in (17') is attractive on empirical grounds as well as on theoretical ground. Data such as the two German sentences below show that the landing sites for scrambling allowed by the Conservation of C-Command are not merely potential landing sites for a scrambled element, but rather that if an element scrambles to a high structural landing site, it must have moved cyclically through the intermediate potential landing sites. In other words, the scrambled constituent leaves a copy at every intervening maximal projection allowed by (17).

(18) Gestern hat [jeder Professor]$_i$ [jedem Studenten]$_j$
Yesterday has every-NOM professor$_i$ every-DAT student-DAT

seine$_{i,j}$ Dissertation gegeben
his$_{i,j}$-ACC dissertation given
"Yesterday, every professor$_i$ gave every student$_j$ his$_{i,j}$ dissertation."
(Lee & Santorini 1994: 286)

The sentence in (18), with the unscrambled order, is ambiguous between a reading in which the dissertations were written by the professors or the students; either the subject or the indirect object is available as a binder for the possessive pronoun in the direct object. However, when the direct object is scrambled over the subject, as in (19) below, one of the binding possibilities is eliminated.

(19) Gestern hat seine$_{i,*j}$ Dissertation [jeder Professor]$_i$
Yesterday has his$_{i,*j}$-ACC dissertation every-NOM professor$_i$

[jedem Studenten]$_j$ gegeben
every-DAT student-DAT given
"Yesterday, every professor$_i$ gave every student$_j$ his$_{i,*j}$ dissertation."
(Lee & Santorini 1994: 286)

When the direct object containing the possessive is scrambled across the subject, it can reconstruct to be bound by the subject (as I discuss at length in Chapt. 8; cf. references there and the discussion of these examples in Lee & Santorini 1994). However, even though the direct object reconstructs to a position below the subject in (19), it reconstructs to a position *above* the indirect object, i.e. an intermediate landing site for scrambling. For this reason, it is clear that the direct object had to scramble to a position below the subject but above the indirect object, leaving a copy there, before it moved on to scramble above the subject. Since "short" scrambling below the subject never reconstructs, the indirect object can no longer bind the direct object once it has moved through a higher position on its way to moving above the subject. The subject thus binds the possessive in the direct object in a derived scrambling position, and this can only be accounted for under a theory in which scrambling leaves copies at each available landing

site as an element scrambles up the tree. Indeed, this result is predicted by Closest/Shortest Attract, combined with the copy theory of movement.

5.5 Conservation of C-Command and Tree Adjoining Grammars

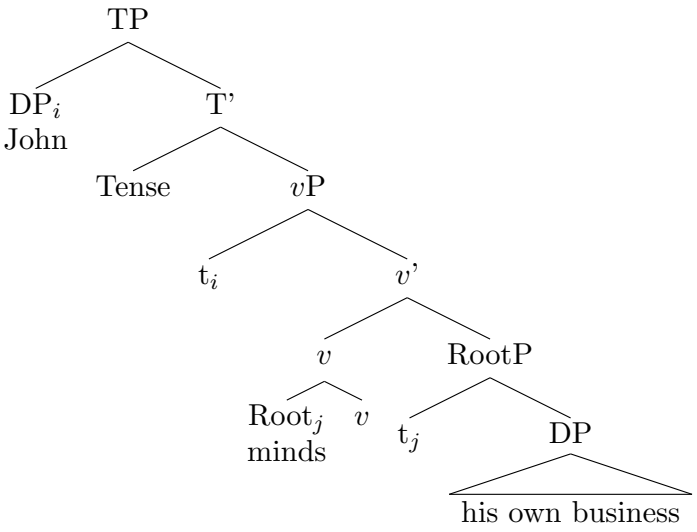
The Conservation of C-Command is the sort of universal that is expected if Tree Adjoining Grammars (TAGs; Joshi, Levy, & Takahashi 1975, Vijay-Shanker & Joshi 1985, Kroch & Joshi 1985, and much subsequent work following these original studies) are a good model of recursion in natural language. In a TAG, adjunction is, by definition, separate from other processes that determine the final structure for a sentence; the fact that adjunction is encoded differently in a TAG is the distinguishing feature of the formalism, so the idea that adjunction is different from other parts of the grammar is axiomatic. However, the question of precisely what sort of object adjoins into what other sort of object is a question outside the realm of basic TAG.

TAG by itself is a remarkably spartan formalism, not making any inherent assumptions about the forms of elementary trees, which are the primitives of any TAG, nor does it say anything about the relationships between different elementary trees other than that they combine with each other. A TAG blindly combines elementary trees in a mathematically elegant and constrained way (*substitution*, and what is frequently referred to as *Joshi-Adjunction*); the work of the syntactician lies in specifying the inventory of elementary trees in a way that captures important linguistic generalizations.

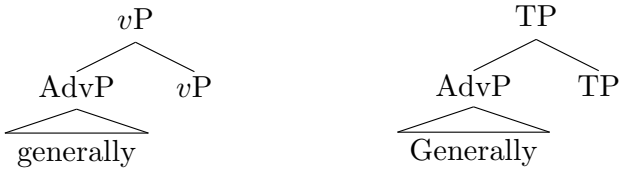
However, the literature on TAG rarely discusses the concept of the relationships between different elementary trees in detail, addressing questions whether there are constraints that apply across the entire elementary tree inventory, or if part of the inventory should be derived by transformations on a smaller number of even more basic trees (preceding the substitution and adjunction operations of TAG proper). Linguists working within a TAG framework usually make a number of implicit assumptions about the relationships between elementary trees, although there have been very few serious attempts to formalize these types of meta-TAG relationships between trees, a notable exception being the "TAG metarules" system of Xia (2001). One very general example of this type of meta-TAG assumption is as follows: every application of TAGs to linguistic research assumes that the set of auxiliary trees (trees which adjoin into other elementary trees) have root nodes which necessarily have the same label as some node in a non-auxiliary tree in the tree inventory; if they did not, they wouldn't ever be able to adjoin anywhere and would be of very little use.

To take a more specific example, every adverbial auxiliary tree, i.e. a tree containing an adverbial frontier node (or in a lexicalized TAG, a tree associated with a single terminal which is an adverb), must have a root node with a label that corresponds to some type of node that adverbials actually adjoin to. For instance, given the tree in (20) below, the adverb *generally* must be associated with auxiliary trees rooted in *v*P or TP as in (21), allowing the derived trees in (22).

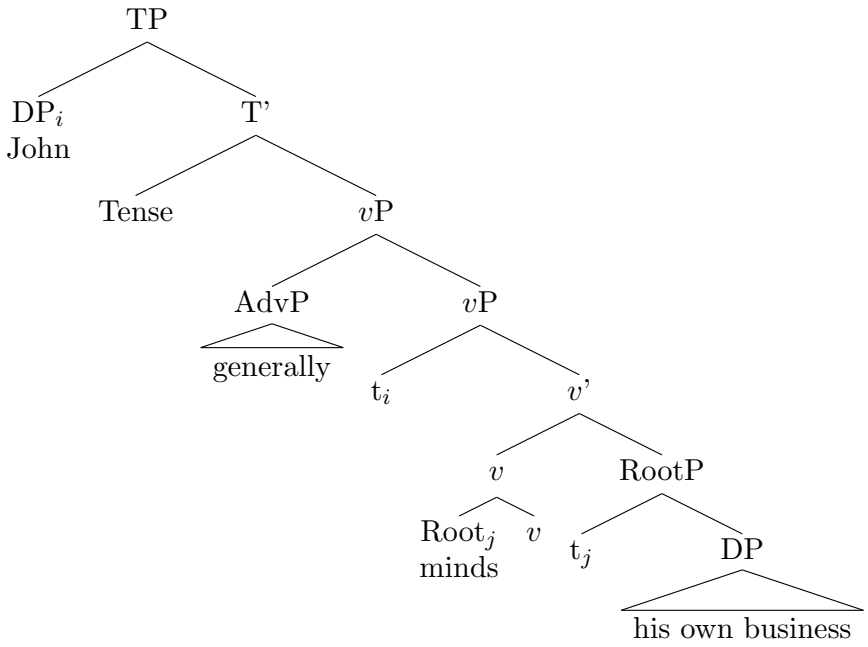
(20)

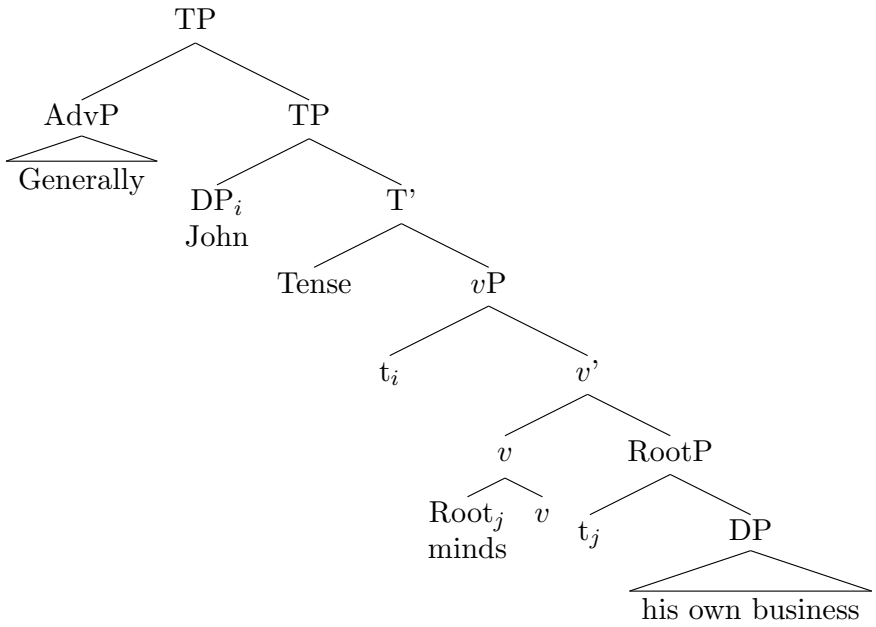

TP
DP$_i$
John
T'
Tense
vP
t$_i$
v'
v
RootP
Root$_j$
minds
v
t$_j$
DP
his own business

(21)


vP
AdvP
vP
generally
TP
AdvP
TP
Generally

(22)


TP
DP$_i$
John
T'
Tense
vP
AdvP
vP
generally
t$_i$
v'
v
RootP
Root$_j$
minds
v
t$_j$
DP
his own business



If *generally* were only associated with auxiliary trees rooted in DP, ungrammaticality would result, even though such a TAG is perfectly imaginable. This is actually a more fine-grained meta-rule/constraint than the one I mentioned above, that auxiliary trees be able to adjoin into *some* elementary tree: instead, we now have a rule across the auxiliary tree inventory stating that adverbial trees must be rooted in certain nodes and not others, namely, nodes associated with the extended verbal projection of a clause (and then there are subsets of adverbials, with manner adverbs only adjoining to *v*P, etc.).

Conservation of C-Command is a simple extension of the TAG metarule that constrains adverbials to be rooted in certain nodes, e.g. vP and TP (also depending on the scope of the adverb, of course). Since adjunction is a primitive operation in TAG, it makes sense that there is a set of metarrules that apply just to auxiliary trees, which are the building blocks of adjunction structures. But in the case of scrambling, instead of a constraint on adverbials, the Conservation of C-Command is a constraint on the

elementary tree inventory which generates auxiliary trees from non-auxiliary elementary trees in the way described in (23) below. These auxiliary trees are generated from elements in the non-auxiliary elementary tree that may potentially scramble (e.g. objects), and their root nodes are derived from the positions in the non-auxiliary elementary tree which they may scramble to. Following Frank (2002), all (non-adjunction) movements occur in elementary trees,[54] including the major (Kaynian) XP-movments required to derive head-final strings as described above. Once these tree-internal movements are complete, the scrambling auxiliary trees are derived by the following algorithm:

(23) Definition: **Scrambling (TAG) Metarule**

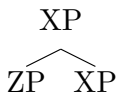
1. Each head node colors the (XP) nodes in its c-command domain with a unique color.[55]
2. Adjunction structures (auxiliary trees) are derived for each object (or other scrambling element, depending on the language), ZP, from the

[54] In fact, movement in this view is actually another type of TAG metarrule, since it takes place derivationally before TAG-proper, which only consists of the adjunction and substitution operations. I assume that Frank's (2002) movement, which is internal to the elementary trees, precedes other metarules such as the Scrambling Metarule, since these movements apply before any consideration of the other trees in the elementary tree inventory takes place (and there is no reason to consider other trees until the time has come for adjunction and substitution). However, this is not really a necessary assumption: it's possible that all metarules apply at all levels of the TAG derivation cyclically, and although additional cycles usually have no further effect on the derivation, applications of the Scrambling Metarule after elementary-tree-internal head-and XP-movement will have an additional effect. In any case, I leave the investigation of this detail of TAG derivations for further research.

[55] This operation was originally suggested to me by Aravind Joshi (p.c.) as part of a way to formalize head-movement in a TAG.

maximal (XP) nodes that dominate ZP and are not a different color from ZP; i.e. an auxiliary tree is generated if color(ZP) ∈ color(XP).

3. The auxiliary trees are of the form:

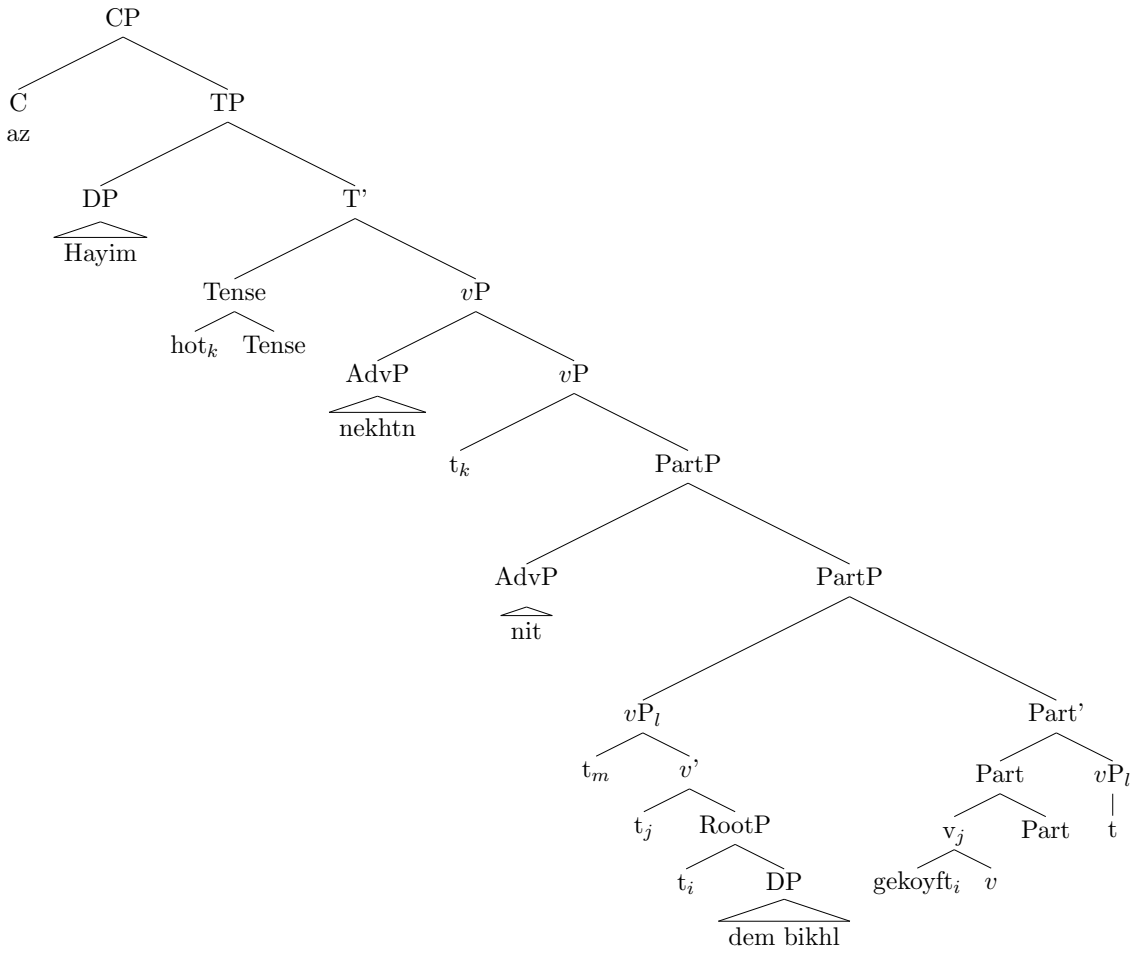


To give an example of the TAG Scrambling Metarule, consider the Yiddish example in (24), repeated from Chapter 2 example (9) above, with the elementary tree shown in (25) after internal movements have taken place (following Frank 2002).[56] (The details of the *v*P-to-Spec(PartP) movement analysis of OV order in this Yiddish clause are discussed in detail with reference to examples 25-28 in the following chapter. For the present purposes, just assume this movement has already taken place.)

(24) … az Hayim hot nekhtn nit dem bikhl gekoyft.
that Hayim has yesterday not the book bought

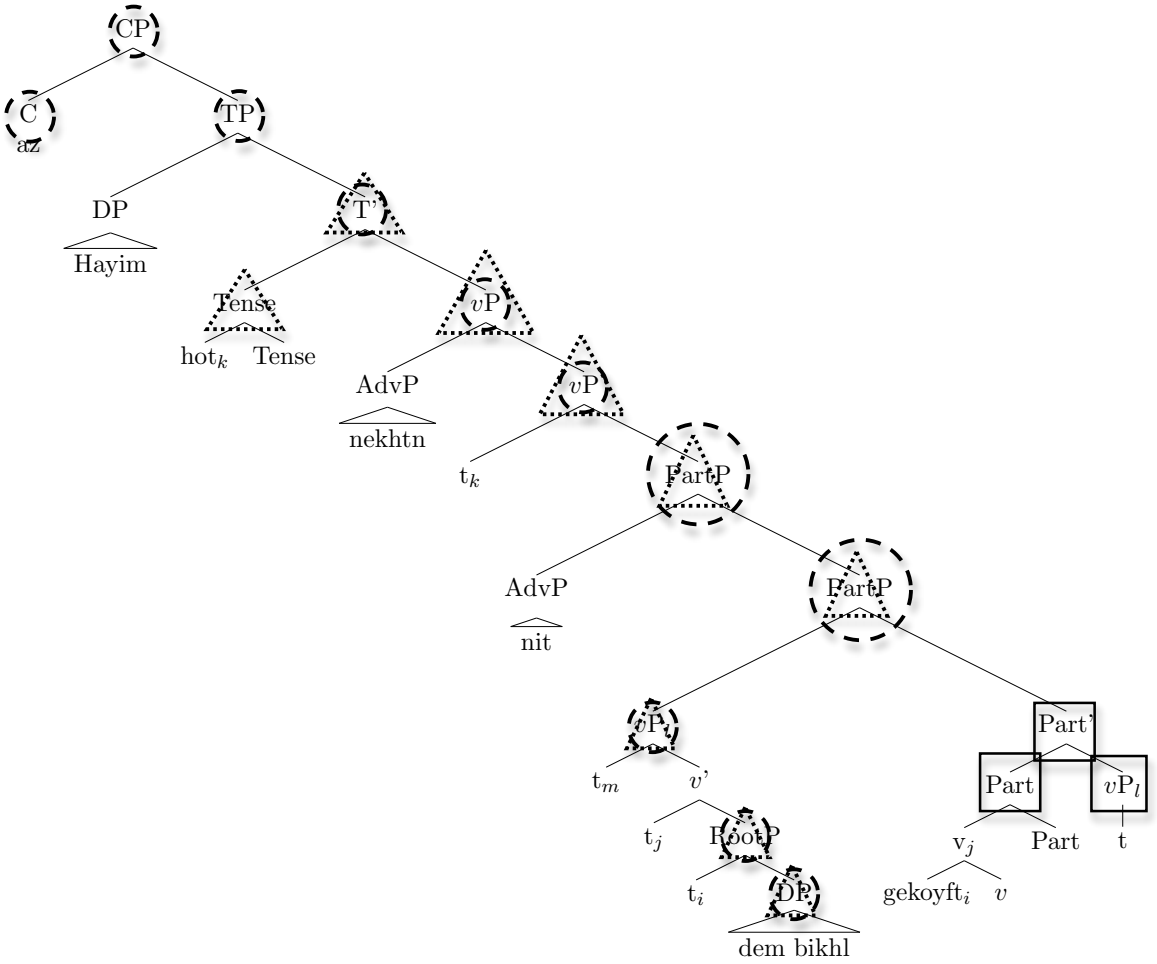
[56] Actually, (25) is not the elementary tree, strictly speaking, since some substitutions have already been made and two adverbials have already been adjoined into the tree. This is done for expository purposes only, to show where the subject has landed and where the object might potentially scramble to, and does not bear on the derivation proposed in this section.

(25)

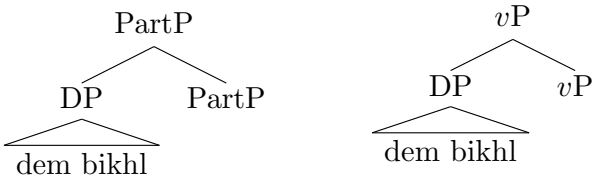


The metarule progresses as follows:

1.



2-3.



The final step of the derivation of scrambling in a TAG is for copies of the scrambled element (in its final position as determined by the scope that the speaker desires), which the scrambled element c-commands to be deleted. While a detailed discussion of PF deletion in TAG is beyond the scope of this dissertation, I will merely note here that

some such notion must be introduced into TAG if the insights of the copy theory of movement are to be retained in TAG-based syntactic analyses.

To conclude this section, I have shown that the Conservation of C-Command is a constraint (or rule) of an expected type under TAG. First, the Conservation of C-Command treats adjunction as different from other syntactic operations, and all TAG analyses predict that this should be a universal feature of natural language. Second, there is a natural statement of the constraint in TAGs which is a subtype of a necessary metarule for any TAG syntax. And lastly, the metarule in its most general form produces copies of the scrambled element: the TAG formalism thus predicts that scrambling should be cyclic (up to the point in the tree where the scrambled element finally lands on the surface), just as I have empirically defended earlier in the discussion of examples (18) and (19) in §5.3.

5.6 Summary

The preceding chapter was intended to show only three things. First, the chapter showed that a highly constrained version of Kayne's antisymmetric phrase structure is possible, and secondly, that it is applicable to the languages of OV Germanic. Thirdly, the chapter showed that the antisymmetric approach to head-finality allows for a restatement of the GHC in purely hierarchical terms, as the Conservation of C-Command, a constraint on the computation of adjunction.

Chapter 6

Results and Predictions of the Antisymmetric Approach to Scrambling

6.1 Introduction

This chapter hopefully picks up where the last one left off: while the last chapter has primarily theoretical results, this chapter shows that these theoretical gains have empirical teeth. The first section below shows that the Conservation of C-Command in fact covers the data it was formulated in order to cover,[57] the typology of scrambling languages I have already discussed. The second section shows that the approach to OV languages developed in the previous chapter, in combination with scrambling and the Conservation of C-Command, generates all of the observed patterns of the West Germanic Verb (Projection) Raising construction, while excluding logically possible but empirically ungrammatical patterns. Furthermore, no additional theoretical machinery is necessary to achieve this result. Rather than being a strange and difficult to capture phenomenon, as it is usually viewed, the Verb (Projection) Raising construction turns out

[57] It would be rather sad if that weren't true, though it wouldn't be the first time in linguistics.

to be an entirely expected part of Germanic clausal syntax and is explainable in terms of operations which are all independently attested elsewhere in Germanic.

6.2 The Typology of Scrambling Languages, Revisited

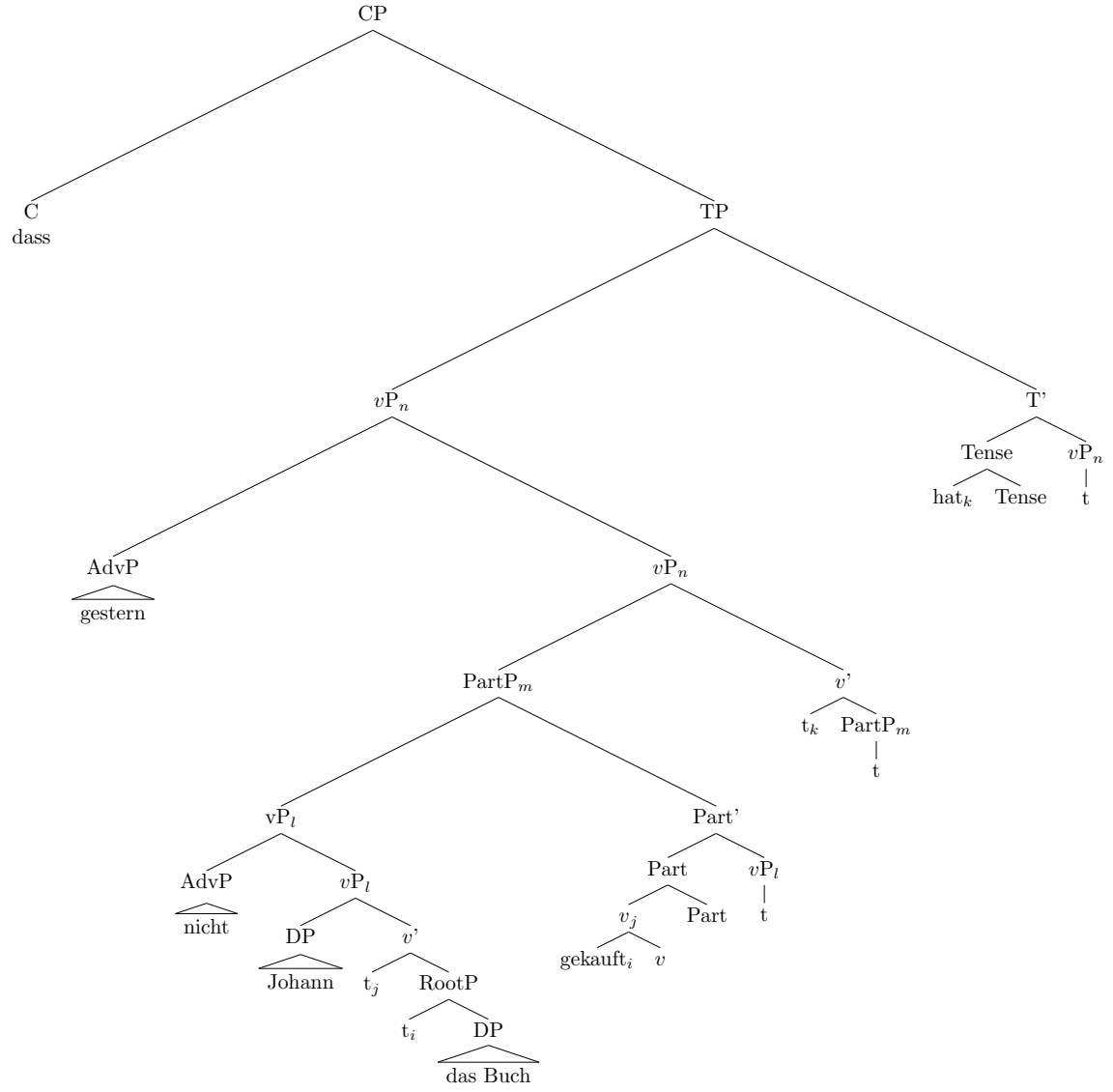
Combining the Conservation of C-Command, the analysis of scrambling as adjunction, and the "roll-up" antisymmetric analysis of head-final projections, it is possible to derive all of the cross-linguistic GHC effects I described in the preceding chapters. To illustrate, I will continue with the case of German, where scrambling potentially targets all maximal projections c-commanded by Comp (as always, given the appropriate prosodic and information structural conditions). The now familiar German scrambling paradigm is below.

(1) a. Ich weiß daß Johann gestern nicht **das Buch** gekauft hat.
I know that Johann yesterday not **the book** bought has.
"I know that John didn't buy the book yesterday."
b. Ich weiß daß Johann gestern **das Buch** nicht gekauft hat.
c. Ich weiß daß Johann **das Buch** gestern nicht gekauft hat.
d. Ich weiß daß **das Buch** Johann gestern nicht gekauft hat.

e. * Ich weiß **das Buch** daß Johann gestern nicht gekauft hat.
(with contrastive, or at least, narrow focus on *das Buch* in "a", but not in examples "b" through "e")

Following the antisymmetric approach to German above, the structure for the subordinate *daß*-clause in (1) after head-movement and the EPP-driven vP/PartP roll-up movement, but before scrambling, is as shown below:[58]

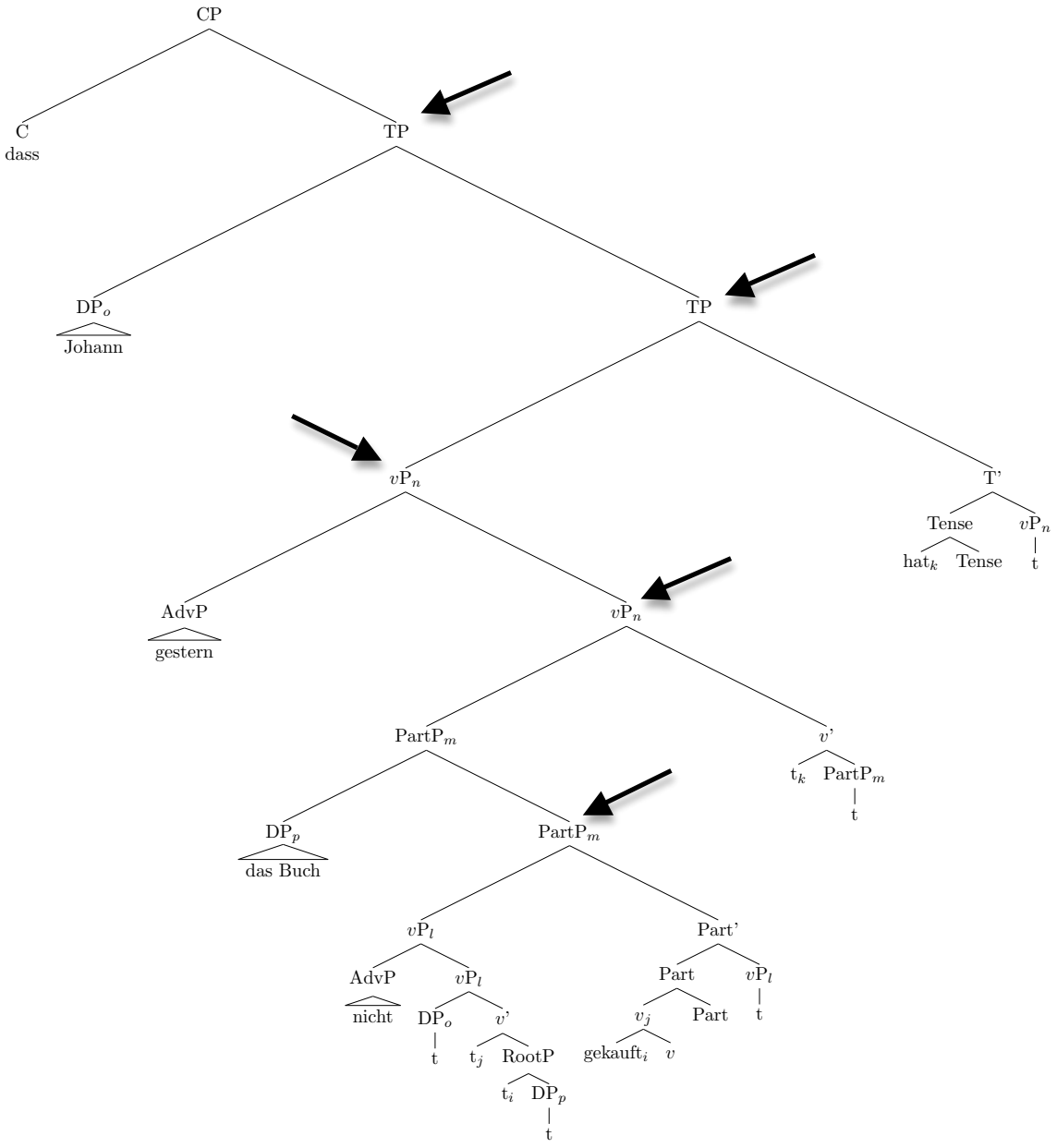
(2)



[58] I assume for the sake of argument that the negative adverb *nicht* ("not") is adjoined at the lowest possible position that still allows it to have scope over the event (including the event's Agent), the lower vP, and that *gestern* ("yesterday") is adjoined at a higher position. However, nothing in this analysis depends on the details of their adjunction sites, and of course, I would expect the adjunction sites to vary with different scope readings (these may not be distinguishable in this particular example).

According to the Conservation of C-Command, the large amount of pied-piping in German (as shown above) creates a variety of possible landing sites available for scrambling. The EPP on Tense plus the cyclic pied-piping of all of the maximal projections along the spine of the clause results in a structure in which the constituents within the lowest vP are c-commanded by only a single head: C. In most cases, depending on information structure, prosody, and the definiteness of the arguments (see Diesing 1992, Diesing 1997, *inter alia*), both the subject and object scramble out of the vP at this point. Taking the case of object scrambling, the only attachment site that is impossible for a scrambled object in German is CP: adjoining at CP would subtract the c-command relation holding between C and the object, resulting in the ungrammatical sentence in 18e. However, the object can adjoin at any lower maximal projection because the c-command relations have already been rearranged by feature-driven A-movement with pied-piping. In the tree below, I show the same tree after the subject has scrambled to TP and the object has scrambled to the position between the two adverbs, shown above in (18b). The arrows indicate other possible scrambling sites for the object.

(3)



In the tree above, the object has scrambled to adjoin to PartP, deriving the string shown in (18b). If the object continued to scramble up the tree and adjoin to all of the sites marked with arrows, it would ultimately derive the sentence in (18d), while also deriving the other grammatical orders in (18) along the way. If the object scrambled to the (18d)

position, leaving a copy at every one of the intermediate adjunction sites marked with arrows, all of the copies would still be c-commanded by Comp at the end of the derivation (and throughout). Thus, scrambling is an optional process, from the point of view of the narrow syntax, but it is a well-constrained one, local to a certain c-command domain.

Yiddish, as I described in Chapter 3, began with the Tense-final system of German and changed one parameter during its history to become Tense-medial, trapping scrambling to positions below Tense; this change must now be understood in a new way, in light of the analysis of German presented above. Under the broadly Kaynian analysis of head-finality, this change can no longer be described as a change in the "head parameter" setting for Tense. Instead, the difference between German/Early Yiddish and modern Yiddish must be in the Tense Probe's EPP feature, along the same lines that Biberauer & Roberts (2005) argue for historical English. Specifically, "Tense-final" languages have [+ pied-piping] associated with Tense's EPP feature, and Yiddish changed to have a [- pied-piping] feature on Tense. Tense-final languages always pied-pipe *at least* one level of structure with the subject to Sepc(TP), i.e. they pied-pipe at least the most embedded vP to Spec(TP). Yiddish changed its pied-piping on Tense over the course of its history, and so no longer pied-pipes this lowest vP to Spec(TP); the introduction of this modern Yiddish non-pied-piping Tense head into the speech community led to the changes described in Chapter 3 and in Santorini (1992), (1993a).

However, the other verbal heads were not all affected by this change. If other heads along the clausal spine remain [+ pied-piping], as in German, then the change in

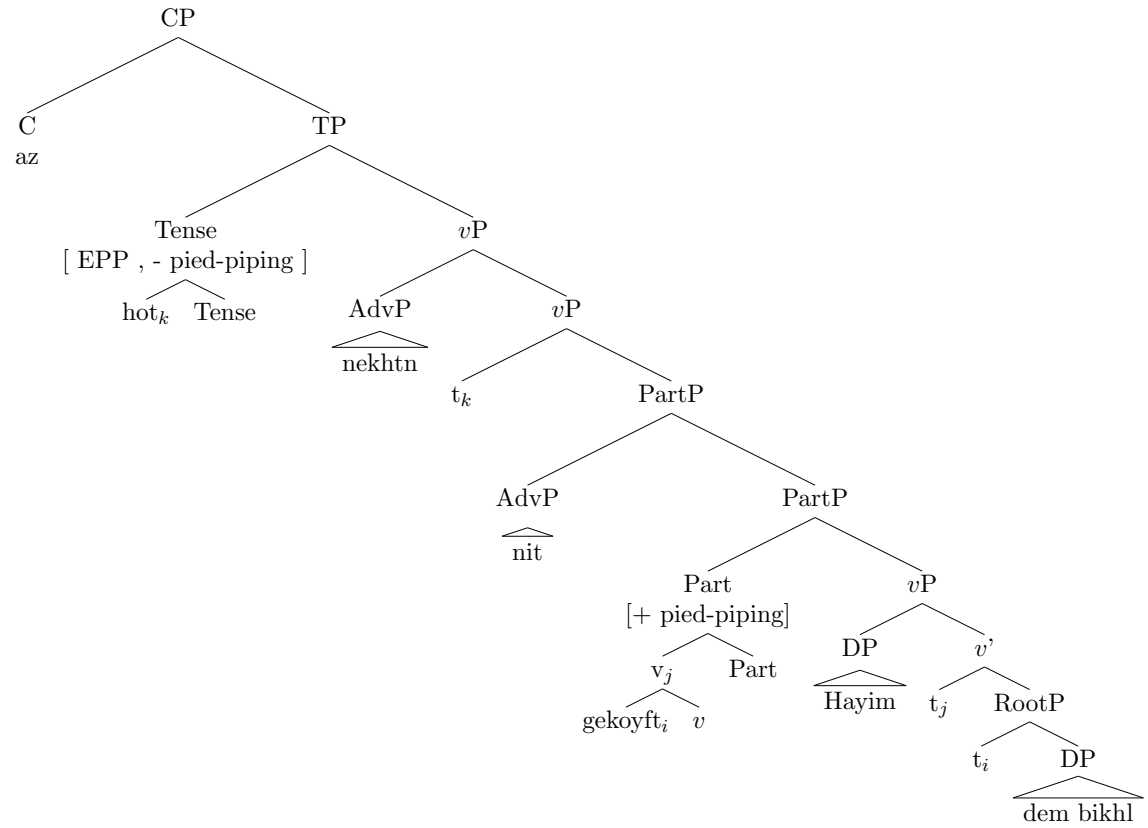
just the [pied-piping] feature on Tense results in the "SIOV" syntax of modern Yiddish. To illustrate, take the familiar subordinate clauses below, and the following pre-XP-movement tree.

(4) a. Ikh trakht az Hayim hot **dem bikhl** nekhtn nit gekoyft.
I think that Hayim has the book-DIM yesterday not bought
"I think that Hayim didn't buy the book yesterday."

b. Ikh trakht az Hayim hot nekhtn **dem bikhl** nit gekoyft.
c. Ikh trakht az Hayim hot nekhtn nit **DEM BIKHL** gekoyft.
(constrastive stress on *dem bihkl*)

d. * Ikh trakht az Hayim **dem bikhl** hot nekhtn nit gekoyft.
e. * Ikh trakht az **dem bikhl** Hayim hot nekhtn nit gekoyft

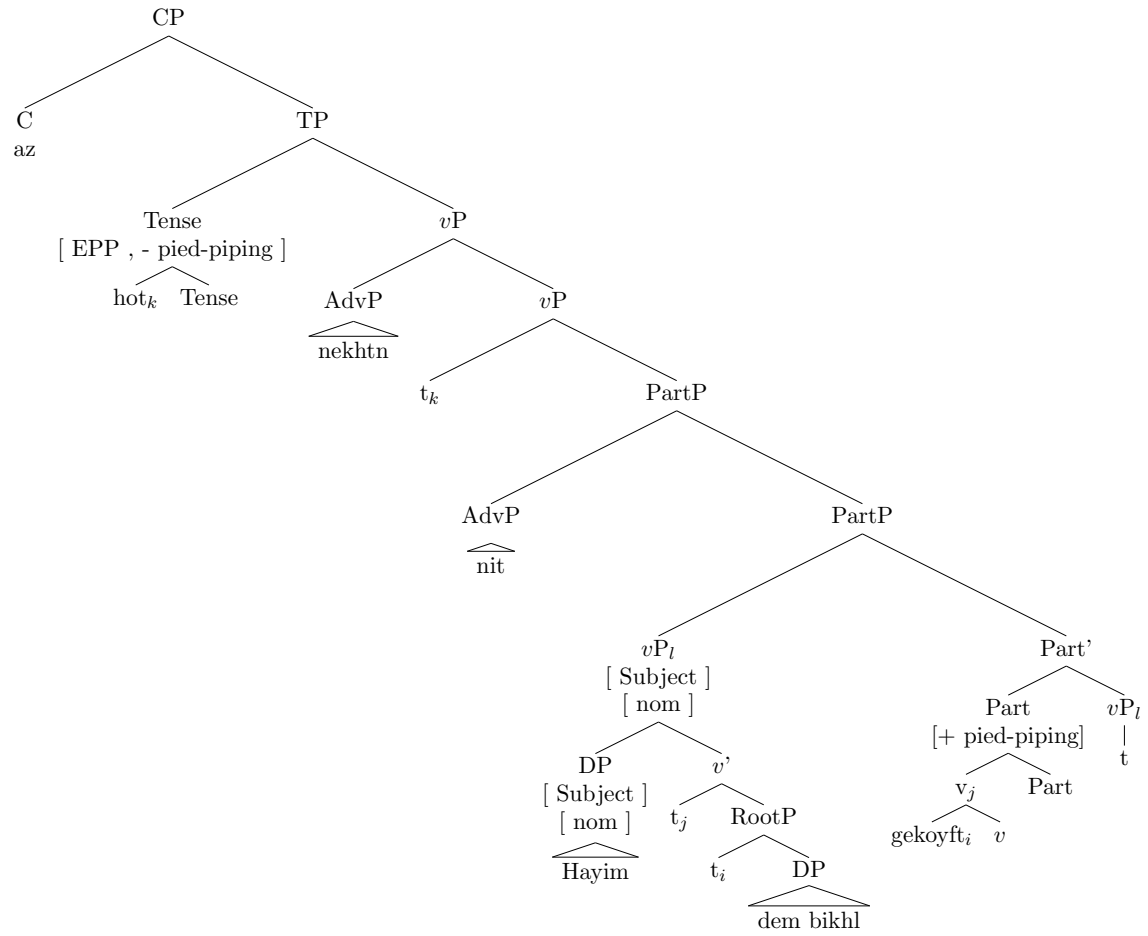
(5)



Tense, a Probe by virtue of the EPP, looks down the tree with a [- pied-piping] feature in modern Yiddish. When it reaches the most embedded *v*P, with the subject in Spec(vP), the [- pied-piping] feature would normally copy onto the head immediately dominating the lowest vP, the Part head in the trees above. This copying would cause the Tense EPP's [pied-piping] feature to be satisfied by deleting any percolated features from the subject in Spec(vP) to its maximal projection, vP. However, this derivation is not available in the case of Yiddish: if the head that the EPP's [- pied-piping] feature is supposed to copy onto (the Part head, in this case) has a conflicting pied-piping feature [+ pied-piping], as I'm suggesting for Yiddish, the head will be specified for two conflicting feature values in violation of the theorem in (11) above. With this combination of a [- pied-piping] Tense-EPP and a [+ pied-piping] Part, the derivation will crash unless the [- pied-piping] feature can be satisfied by deleting percolated features at some other stage of the derivation where the feature-conflict does not arise.

It does occur at another stage of the derivation, one step of the derivation later, producing the SIOV pattern characteristic of Yiddish, Middle English (in part), and the Kru languages (Koopman 1984). The lowest vP moves to Spec(PartP), satisfying Part's [+ pied-piping] feature, which percolates the subject's features to the maximal vP, as shown below:

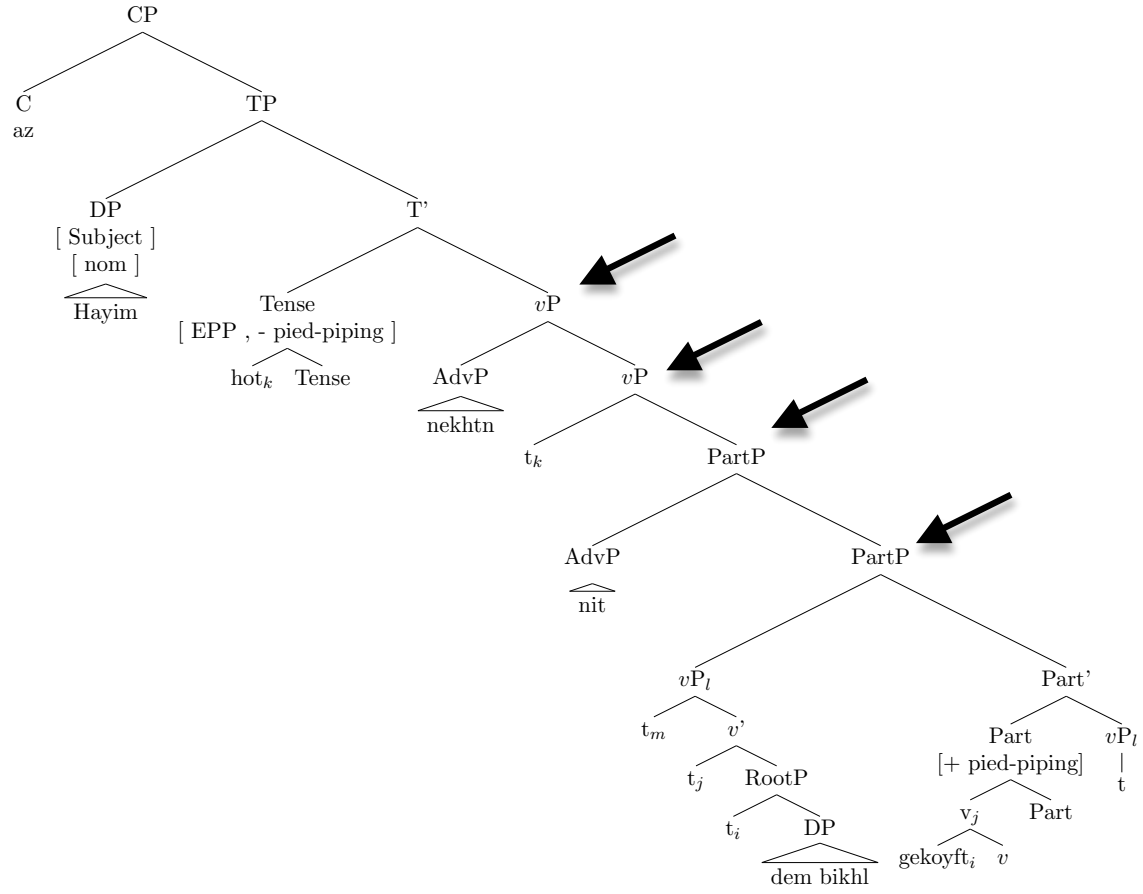
(6)



At this point in the derivation, Tense's EPP [- pied-piping] feature can be satisfied. It operates on the vP dominating *Hayim* and deletes the percolated subject features, leaving only the subject to be attracted to Spec(TP).[59] The derivation is completed as shown in (7) below.

[59] Whether the Tense Probe's [- pied-piping] feature is copied onto the trace of the little-v above *nit* and deletes the percolated features there, or whether this deletion occurs at the complex Tense head itself is a technical detail with no consequences that I can see (some would consider this detail to be *nit*-picking – Jonathan Gress Wright, p.c.). I also assume that the adverbs do not intervene in any way vis-a-vis the ability of a head to cause feature-percolation in its complement; indeed, nothing is lost if the adverb are simply taken to adjoin later in the derivation.

(7)



The adjunction sites that the Conservation of C-Command makes available for scrambling are indicated by arrows in the tree above, and these correctly generate the grammatical orders in (4) above while excluding the ungrammatical orders in (4). Scrambling in modern Yiddish is correctly restricted to sites which are to the right of Tense stringwise, but the constraint is a purely hierarchical one. The difference then between Early Yiddish and German on the one hand, and modern Yiddish on the other, is also purely hierarchical. The change in the stringwise position of the tensed verb in Yiddish, described in Santorini (1992, 1993a), and it's affects on the landing sites available for scrambling, as I described in Chapter 3, resulted from the introduction of a

new [EPP, - pied-piping] variant of the Tense head into the Yiddish speech community.[60] In this way, the competition between the "Tense-final" Yiddish phrase structure and the "Tense-medial/initial" phrase structure, discussed in Chapter 3 above and in Santorini (1992, 1993), was competition between the two variants of the Tense head. This result is entirely in line with the hypothesis presented in Kroch (1994) that grammar competition can always be described in terms of competition between functionally equivalent syntactic heads carrying different morphosyntactic feature specifications.

This way of stating the variation, in terms of grammar competition, may appear strange at first to many readers, mostly because much of the discussion of competing grammars in the literature has focused on competing settings of the headedness parameter. Indeed, some researchers (e.g. Hinterhölzl 2009) have suggested that an antisymmetric account of OV languages makes the notion of grammar competition obsolete, by the very fact that it relates OV and VO languages by a difference in XP movement rather than in a difference in headedness. However, this type of argument reveals a serious misunderstanding of the notion of grammar competition as first set out in Kroch (1989) and developed in much subsequent work. "Grammar competition"

---

[60] There is one wrinkle in this story, stemming from the theorem in (11) above. In order for there not to be a feature-conflict between Tense and the first little-v head below Tense, when Tense changes to [- pied-piping], that little-v must either change to [- pied-piping], change to be underspecified for [pied-piping], or already be either [- pied-piping] or underspecified for [pied-piping]. It may be enough for there to be a version of that head in speakers' inventories which is either [- pied-piping] or underspecified for [pied-piping], as some little-v heads are in languages with the West Germanic verb-raising construction (e.g. Swiss German, Dutch, West Flemish, and indeed, Early Yiddish), as I will argue below. I will leave the full solution to this problem as an open question for the time being, but I will note here that if the change to Tense-medial is dependent on there already being a [- pied-piping] little-v in the language, this implies that there is a close relationship between the West Germanic verb-raising construction and a language's ability to change from Tense-final to Tense-medial, an hypothesis that might prove a fruitful direction for further research.

simply refers to a situation in which two (or more) functionally equivalent linguistic variants unstably coexist in a population. Because humans are capable of learning multiple forms for one function, even if they are incompatible with a single grammatical system (i.e. bilingualism or bidialectalism), any time this type of unstable variation exists in a population, it will also exist within the performance of individuals. The two grammars only "compete" in the sense that a speaker can only produce one of the two at a given time; in other words, in a single utterance containing a single appropriate context for the two variants, it's only possible to produce one at a time, and so they vie for linguistic space/time, in a sense. If one of the two variants is favored for some reason (i.e. "fitter" in an evolutionary sense) in terms of transmission across generations ("reproduction") and/or social resources, then the system is unstable, and it will naturally increase in frequency in the speech of individuals and the community at the expense of the other variant. Thus, "grammar competition" is merely evolutionary dynamics, but applied to linguistic variables rather than genotypes (see Nowak 2006, who naturally interprets "grammar competition" as just another type of evolving, and therefore unstable, biological system).

The question of whether there are competing grammars in a population or an individual is thus entirely orthogonal to the question of the best theoretical understanding of a syntactic variant. In the Yiddish case, describing the difference between the Early Yiddish grammar and the Modern Yiddish grammar in terms of vP/VP-fronting rather than in terms of headedness is an entirely independent question of whether grammar competition exists. Under either analysis, the following holds true: there was some older

syntactic language, it came into unstable variation with a new syntactic variant, and over centuries, the new variant was used more and transmitted more effectively until the old syntactic variant disappeared from Yiddish.

To complete the discussion of antisymmetry and scrambling, the derivation of complementizer-final languages like Japanese, Korean, or Amharic progresses almost in the same way as the one for German, and is similar to the German derivation in more ways than one. As in Tense-final languages like German, the heads along the clausal spine are specified for [+pied-piping], triggering a complete roll-up of the structure as we've already seen above. The pied-piping mechanism I've described thus far is sufficient to derive a surface tensed-verb-final surface structure, but in order to derive the complementizer-final surface pattern of Japanese and Korean one additional movement with one additional level of pied-piping must take place; somehow, the TP must move to surface to the left of the complementizer position in these languages. The trigger for pied-piping has an obvious source: since all functional heads in Japanese[61] (and other such relentlessly head-final languages) appear finally in their phrases, I propose that functional heads in Japanese are all specified [+ pied-piping] (at least by default). By extension, Comp is also [+ pied-piping] in Japanese.[62] However, this does not solve the problem entirely: the EPP and nominative case-checking have already been satisfied by the subject's movement to Spec(TP), which also pied-piped the various verbal projections to that position. Therefore, some additional movement trigger must exist to cause the TP

---

[61] DP appears to be the only possible exception in Japanese, though there isn't clear consensus on the structure of the Japanese DP in the literature.

[62] Or the Force head, presumably, in the "Split-C" phrase structure of Rizzi (1997), Roberts (2005), among others.

to move past C, and such an analysis is only plausible to the extent that there is some other movement that has independent motivation; otherwise, the antisymmetric analysis of complementizer-final languages becomes just an ad hoc placeholder for some better analysis in the future.

Fortunately, there is such an independently attested motivation for movement: the verb-second (V2) phenomenon of (at least) continental West Germanic, North Germanic, and Kashmiri (Bhatt 1999). While the precise nature of the V2 constraint is not currently well-understood, there is has been a general consensus since (roughly) den Besten (1983) and Haider & Prinzhorn (1986) that it consists of a requirement to fulfill two conditions: first, that the finite verb move to C (unless that position is occupied by a complementizer), and secondly, that some element merge in the Spec(CP) position creating a spec-head configuration with the element in C.[63] By default, the Spec(CP) position is filled by the thematic subject of the clause, though it is also frequently filled by other constituents depending on information structural considerations and whether a thematic subject is available. I propose that complementizer-final languages are simply another type of V2 language (even though "verb-second" is hardly an appropriate description in this case) which share at least the second V2 condition with languages like German; they probably also share the condition of verb-movement to C unless a complementizer merges in C, making the parallel with German complete, though T-to-C movement will be string-vacuous in every case.

---

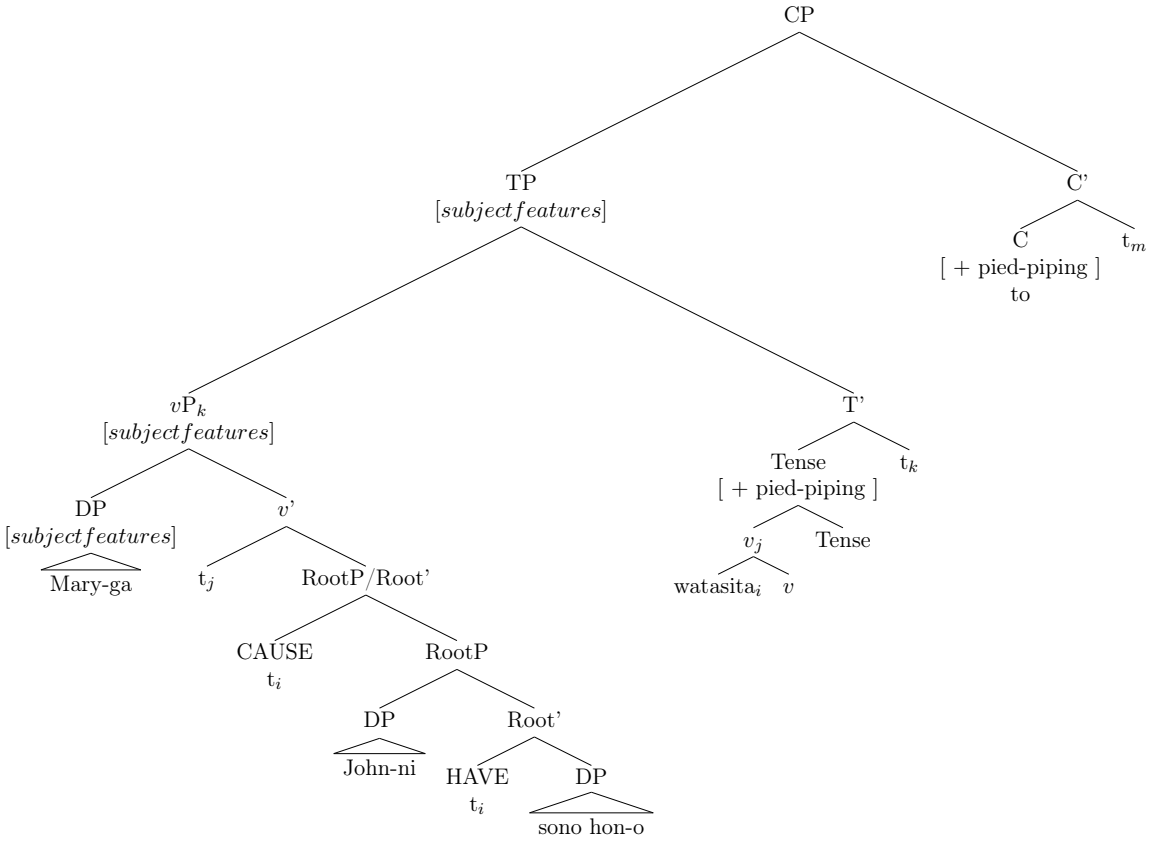
[63] At least, this is the consensus for "CP-V2" languages such as German, Dutch, and mainland North Germanic. It may also be true for Yiddish and Icelandic matrix clauses, but the question of whether the V2 requirement is fulfilled at the IP/TP level in those two languages (cf. Diesing 1990, Thráinsson 1986, Thráinsson 2007) is still a matter of debate.

The main difference between Japanese and German under this approach is that C in Japanese is [+pied-piping], which causes the subject's features in Spec(TP) to percolate one more level to TP itself. Additionally, I also need to hypothesize one further difference between German and complementizer-final languages. In the former, the subject, object, etc., or various sub-constituents of those elements may move to Spec(CP) to satisfy the V2 constraint, depending on the information structure of the clause. This indicates that the V2 constraint is not satisfied by C attracting some purely formal feature of the constituent like Case or whichever features are relevant for the EPP on Tense. However, in languages like Japanese, the part of the V2 constraint which forces Spec(CP) to be filled must have become more grammaticized than it is in German, and for that reason the V2 constraint is sensitive to the same, purely syntactic features that the EPP on Tense was. In this way, the closest element bearing subject features must be attracted to Spec(CP) in a complementizer-final language, which is the TP if C has caused the subject's features to copy onto TP by virtue of a [+ pied-piping] specification. Thus, in a sentence like the one below, the subject's features have been copied to the embedded TP, causing TP to front to Spec(CP) as shown in the tree for the subordinate clause below the example sentence. (Of course, the complete derivation of the matrix clause will include further movement of the complement clause with the matrix subject to the matrix Spec(TP) and then matrix Spec(CP)).[64]

---

[64] The ditransitive structure in the tree below follows Harley (2002) in general, though nothing in the present discussion depends on these details. The "CAUSE" and "HAVE" labels are only meant to indicate where the decomposed heads making up the ditransitive structure originally were; of course, these heads have already moved and incorporated with a number of functional heads at this point in the derivation, as indicated by the traces.

(8) Bill-ga Mary-ga John-ni sono hon-o watasita to itta (koto)
Bill-NOM Mary-NOM John-to that book-ACC handed that said (fact)
"Bill said that Mary handed that book to John."
(Saito & Fukui 1998: 443-444)



As is clear from the tree above, the continued roll-up of the structure with the movement of TP to Spec(CP) takes all of TP's constituents out of the c-command domain of C, and so the Conservation of C-Command does not prevent "long" scrambling out of the clause in complementizer-final languages.

German, on the other hand, has not undergone the change from default subject-movement to Spec(CP) to obligatory subject-(feature)-movement to Spec(CP), nor is German's C [+pied-piping], so much smaller constituents fill Spec(CP) in that language

and scrambling is bounded by C. Indeed, there is some evidence to suggest that this was actually an historical development in Japanese, and that the pre-modern Japanese system was actually even more like the German system. In pre-modern Japanese, some C elements could appear in a non-final position, possibly indicating that in the earlier system, smaller constituents than TP could be fronted to Spec(CP) to create the appropriate configuration (Whitman 2001: 96-97).

6.3 The West Germanic Verb (Projection) Raising Construction

6.3.1 The Verb Raising Construction

The Conservation of C-Command, coupled with the antisymmetric pied-piping approach to OV languages presented in the last chapter, predicts the existence of the West Germanic verb (projection) raising construction and provides a clean account of the surface patterns which that class of construction produces across West Germanic. As stated above, pied-piping is driven by the specification of [pied-piping] features on the functional heads, and as the subject of a clause is attracted up the structure by the Tense probe, different specifications for [pied-piping] on heads along the clausal spine determine how much structure is pied-piped with the subject to Spec(TP). What if the

various verbal heads on the clausal spine bore different feature-values for [pied-piping], and/or showed variation in their settings for this parameter? This is precisely the situation in the varieties of West Germanic which show the "Verb (Projection) Raising" construction, either optionally or obligatorily with specific combinations of auxiliaries and forms of lexical verbs (Evers 1975, Zaenen 1979, Haegeman & van Riemsdijk 1986, Kroch & Santorini 1991, Wurmbrand 2004, 2005 and references therein, *inter alia*). In this section I'll show that the difference between English, German, and verb-raising varieties such as Dutch, Swiss German, and Afrikaans, lies in the extent to which additional structure is pied-piped along with the subject to Spec(TP) and, at times, stranded at the subject's intermediate landing sites on its way to Spec(TP).

Languages like English (and e.g. the North Germanic languages) simply do not pied-pipe additional structure with the subject when it is attracted to Spec(TP). Languages like standard German (and, to a first approximation, Japanese, Korean, Hittite and Amharic) generally pied-pipe the maximum amount of structure as the subject moves to Spec(TP), leading to strict surface head-finality in the verbal domain.[65] Between these

---

[65] The only exception in standard German, to the extent that standard German is anyone's native language, is the IPP ("Infinitivus Pro Participio") Construction in which the *haben* auxiliary and the perfect of a modal show the verb-raising order (see example 36 below). This construction is so-named because the nonfinite auxiliary exceptionally has the morpho(phono)logical form of an infinitive in this context, and it also appears in the fully verb-raising orders (i.e. the patterns which appear fully head-initial) in the West Flemish (33) below and the Early Yiddish (34) below. While I will not give a full account of the morphological alternation in the IPP construction, I do take it to be an instance of verb-raising syntactically. In fact, it's possible that the morphological effect is due to the presence or absence of a local [pied-piping] feature, according to the system I have built here: the effect would be explained if the "participle"-morphology spell-out rule for the modal is only triggered when the modal is string-wise adjacent to a head with [+ pied-piping], and the "infinitive"-morphology spell-out rule is triggered in nonfinite contexts where there is no such adjacent [+ pied-piping] feature.

However, the proviso "standard German" is crucial here: Wurmbrand (2004, 2005) shows that when self-described standard German speakers give judgments on 3-verb clusters, where processing is more difficult and the prescriptions of the standard language are less salient, it is possible to see that these speakers actually use verb-raising orders in a larger set of contexts.

two extremes is a wide amount of variation. Much of this variation can be seen within West Germanic, both across West Germanic dialects and across auxiliary verbs within individual dialects/idiolects. In the system put forth in this dissertation, the variation follows from the way the [pied-piping] features are specified on the heads in the extended verbal projection without any further stipulation or the positing of ad hoc types of extraposition. In addition to predicting the correct set of surface patterns, the approach developed in this section goes beyond that result to make two more fine-grained predictions. First, we predict that if a language has scrambling and verb-raising, then it **must** also have projection-raising as a possibility. Secondly, verb (projection) raising structures involving 3-verb clusters have 6 logically possible orderings, but only 5 of those orderings are actually attested in West Germanic. The analysis I present below predicts exactly those 5 orders and rules out the 6th, purely by the way the formalism works: no ad hoc constraint is necessary. The running theme of this section is not going to be that my analysis is the only possible analysis of West Germanic verb projection raising (see Wurmbrand 2004 and Wurmbrand 2005 for a nearly exhaustive enumeration of the possible theoretical approaches). Rather, I submit that this approach is a cleaner analysis with stronger predictions than many of the other possibilities, and most importantly, it does not need to stipulate any additional theoretical constructs in order to cover the phenomena (unlike e.g. Hinterhölzl 2006, where nearly each new pattern requires the positing of a new projection and a new movement). In this way, the approach here has an unusual property which a truly correct analysis must have: it accounts for the

observed phenomena without adding anything to the theory that you haven't already seen in this thesis.

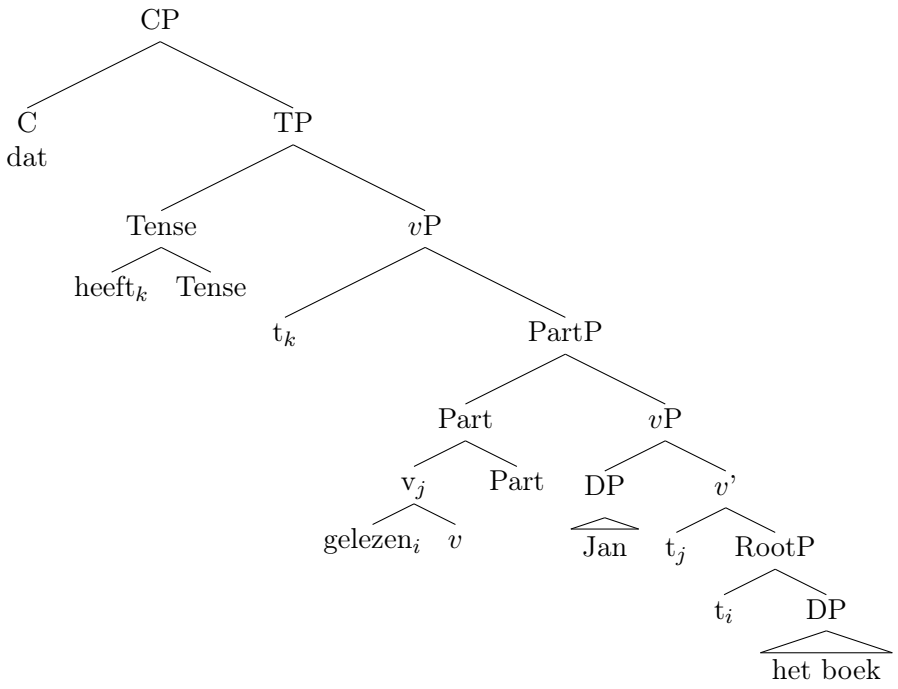
In the antisymmetric view of phrase structure I have been developing here (following in the tradition of Biberauer 2003a), the West Germanic verb-raising construction reflects the underlying order of functional heads on the clausal spine (i.e., in the extended verbal projection), as shown in the Dutch and Afrikaans examples below.

(9) dat Jan het boek heeft gelezen
that Jan the book has read
"…that Jan has read the book"
(Dutch, Wurmbrand 2005: 3)

(10) dat Jan môre kan werk
that Jan tomorrow can work
"…that Jan can work tomorrow"
(Afrikaans, Robbers 1997)

This was the starting point for the derivation of German in the last chapter. In both cases, the derivation begins with only left-headed projections. Head-movement applies, V-to-Part-to-*v* and *v*-to-T, and the intermediate result of the derivation is below.

(11)

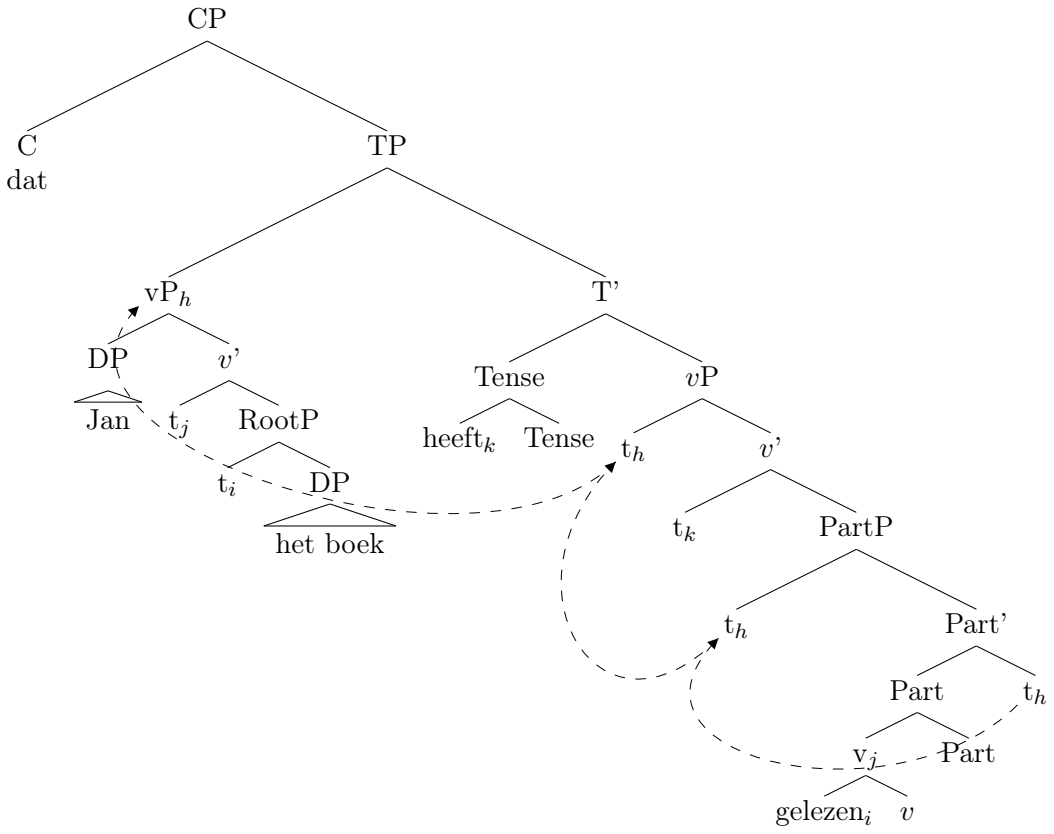


At this stage of the derivation in both German and Dutch, the Tense head probes down the structure to attract the subject, *Jan*. As in all "Tense-final" languages (or "Infl-final" in the classical terminology), Tense carries with it a [+pied-piping] feature which is propagated to the lowest head in the structure above the Goal, in this case *Part*. The [+pied-piping] feature causes the subject's features to percolate up one level of structure to *v*P, ensuring that *v*P will be attracted with the subject to Spec(TP) as the derivation proceeds.

The derivations of Dutch and German diverge at this point. In the Dutch and Afrikaans clauses above, the Tense probe triggers one level of pied-piping, but the little-*v* and Part heads are unmarked for [pied-piping]; they are inert, neither creating nor destroying any feature-percolation that would result in pied-piping. This has a simple result: rather than a "roll-up" of the structure, the lower *v*P moves cyclically through the

available specifier positions along the clausal spine until it reaches Spec(TP). As shown below, the surface structure has left the functional heads in the extended verbal projection to remain in their underlying positions.

(12)



This basic approach to the West Germanic verb-raising construction is a productive one for several reasons. First, it does not add to the theoretical machinery I have already proposed in order to account for head-final surface orders and scrambling under the antisymmetry hypothesis. We paid an initial theoretical price in order to translate the classical head-parameter-based syntax into an asymmetric phrase structure

with pied-piping of large XPs with the subject, though even there I would argue that each piece of old theoretical machinery did not cost more than one piece of new theoretical machinery. But once the pied-piping formalism is in place to derive strictly head-final languages, no additional stipulation is necessary in order to derive verb-raising languages; indeed, the feature system for pied-piping actually predicts the existence of verb-raising languages simply by specifying [± pied-piping] locally on each functional head. In contrast, the classical system does not predict the existence of verb-raising with the same system it uses to derive head-final structure. The classical system actually makes it more difficult to derive verb-raising from the same initial structure that produces surface head-final languages.

Secondly, under the analysis here there is still a common factor tying all of the Tense/Infl-final languages together as a natural class: at least one level of structure, the *v*P immediately dominating the subject, is pied-piped with the subject to Spec(TP). The formal expression of this commonality is the [+ pied-piping] specification of the Tense Probe. There is a single base order for the verbal heads in each of these languages without any ad hoc operations such as "clause union" followed by a supposed morphophonological reorderings of heads. At the same time, we are also able to capture the generalization that the verb-raising order of verbs is the same as the underlying order of verbs in VO languages. Under any of the traditional head-final analyses, the underlying order is head-final, and the fact that some rule of extraposition or of morphological reordering results in the same surface order that VO languages show is a complete coincidence.

6.3.2 The subject's *v*P moves as a constituent

The analysis I have proposed for the West Germanic Verb (Projection) Raising construction, as well as the analysis of the standard German "roll-up" pattern, depends on the movement of the subject's *v*P as a constituent to Spec(TP). In other words, the subject moves with other constituents of the *v*P as a constituent. This does not mean that the lowest *v*P can never be broken up, but rather that it moves as a constituent for some purposes. Under a theory in which subjects in OV Germanic must move to Spec(TP) in order to be licensed (for the EPP to be satisfied and/or for nominative case-assignment), the constituency of the lower vP must necessarily be broken up in order for the derivation to not crash. On the other hand, under the present theory, the constituency of subject and other vP elements is not necessarily disturbed, as long as some other independent process has not taken place (e.g. scrambling of the subject, or topicalization of the subject to Spec(CP) in a matrix clause). So the question at present is: is there any independent evidence that the subject does not always move out of the *v*P? The answer is: yes.

First, there is evidence from the "VP-topicalization" construction in German that the *v*P can move as a constituent with the subject in situ, under certain conditions. The examples below show a topicalized "VP" (probably PartP under the structure assumed here) which has moved to Spec(CP) in a matrix clause as one constituent with the subject. Since the V2 constraint in German prevents the topicalization of more than one

constituent at a time, it is clear that the topicalized material is a single constituent. The subject has not moved to Spec(TP) on its own, nor to AgrSP, which is usually assumed to be higher than TP in the relevant framework. I argue with Haider (1987, 1989, 1993) that VP-topicalization sentences like those below provide evidence that the subject in German has no requirement to move out of its initial *v*P to Spec(TP) or some higher *v*P-external specifier.

(13) (?)[66] Linguisten gespeist haben hier noch nie
Linguists dined have here yet not
"Linguists haven't dined here yet"

(14) (?) Ein talentierter Physiker gelehrt hat hier noch nie
A talented physicist taught has here yet not
"A talented physicist hasn't taught here yet"
(Haider 1989: 193)

As Haider notes, these examples all contain unergative verbs, and so the subjects fronted with the nonfinite verb are unquestionably external subjects. This is assuming that *lehren* "learn" is truly intransitive in 14; of course, if it is really transitive, then the subject is likewise external, and some speakers will also accept the following (a fact that Haider does not address).

(15) ? Ein talentierter Physiker Mechanik gelehrt hat hier noch nie
A talented physicist mechnics taught has here yet not
"A talented physicist hasn't taught mechanics here yet."

[66] Haider (1989: footnote #4) states that the "slight deviance is typical for structures with the external argument, i.e. the NOM-NP, in a preverbal projection. It occurs with ergative subjects as well".

These examples do not mean that the subject never moves out of its base position, of course, but they do mean the following: sometimes the subject does **not** move out, and the subject can be perfectly well licensed without moving alone to Spec(TP). If an entire *v*P or PartP moves to Spec(TP), of course, then the same licensing of case and the EPP can take place without disturbing the constituency that makes the VP-topicalization possible in the sentences above.

Similarly, the placement of manner adverbs shows that the subject does not always leave its *v*P. The paradigm below shows that while the subject (and other constituents) may move over manner adverbs, they need not do so in every circumstance.

(16) a. Ich habe im Hauptbahnhof gewartet
I have in-the train-station waited

als langsam ein schönes Mädchen aus dem Zug ausstieg.
while slowly a-NOM beautiful girl from the train alighted
"I waited at the train station while a beautiful girl got off of the train."

b. ….als ein schönes Mädchen langsam aus dem Zug ausstieg
while a-NOM beautiful girl slowly from the train alighted
(Both orders are grammatical, but "a" is preferred over "b" for this sentence; Beatrice Santorini, p.c.)

(17) Während Hans die Affen beobachtete,…
While Hans the apes observed

a. …gab vorsichtig ein Männchen einem Weibchen einen Kuss.
gave cautiously a-NOM male a-DAT female a-ACC kiss
"While Hans observed the apes, a male cautiously gave a female a kiss."

b. … gab ein Männchen vorsichtig einem Weibchen einen Kuss.
gave a-NOM male cautiously a-DAT fema a-ACC kiss

c. … gab ein Männchen einem Weibchen vorsichtig einen Kuss.
gave a-NOM male a-DAT female cautiously a-ACC kiss

d. * … gab ein Männchen einem Weibchen einen Kuss vorsichtig
gave a-NOM male a-DAT female a-ACC kiss cautiously

In the two sets of examples above, the "a" variant represents the base order and the other variant(s) represent orders in which arguments of the verb have been scrambled past the adverb. The adverbs in the examples, *langsam* ("slowly") and *vorsichtig* ("cautiously"), are low-adjoining manner adverbs; they modify the event alone, and so are standardly assumed to adjoin to the lowest possible projection containing both the predicate and its arguments, the subject's *v*P. Taking these adverbs as diagnostics for the left edge of the lowest vP, it is clear that the subject remains low, inside the vP, in these examples; indeed, this is the preferred position in (16).

Note that the appearance of various constituents of the *v*P between the manner adverbs and the verb (or the verb's position prior to movement to C, in the case of the (14) sentences) is contrary to a claim in Hinterhölzl (2000, 2006). In fact, it shows that Hinterhölzl's "VP-evacuation"-style analysis is trivially wrong. If you carefully construct an appropriate context, not only is it possible to have an object below a manner adverb, but it is actually obligatory (as in 17d). It is also the preferred position in sentences with *v*P-focus; for instance, in the example below, the natural answer to the question "What did he do?" (which implies focus on the event rather than on the object) is the order with the object to the right of the manner adverb.

(18) Er hat sorgfältig die/eine Briefmarke aufgehoben.
He has carfefully the/a stamp up-picked.
"He carefully picked up the/a stamp (as opposed to doing something else)"
(example and judgment from Beatrice Santorini, p.c.)

Hinterhölzl's Kaynian "VP-evacuation" analysis of German depends on sentences like (16) being ungrammatical, which is simply incorrect. He takes the manner adverb to mark the left edge of the lowest *v*P, as I do, and then proposes that the OV/Tense-final order of German is derived by base-generating all of the constituents of the *v*P to the right of the verb and moving them individually to positions on the left. Since manner adverbs can only adjoin low for semantic scope reasons, Hinterhölzl argues that his analysis would receive considerable support if object DPs in German obligatorily surfaced to the left of manner adverbs; this would confirm that the objects have moved to previously unoccupied specifier positions above the *v*P, as he proposes. Unfortunately, this is simply not true. Objects can appear below manner adverbs, and this fact means that they cannot have moved to specifier positions above *v*P in that case. The analysis I have proposed in this dissertation, on the other hand, derives the OV order by moving the remnant *v*P as a whole to the left of the nonfinite verb's position, and so there is no reason that the object and manner adverb can't remain in their original relative positions.

Haeberli (2005) makes similar observations about the existence of a low subject position in German, and he takes this set of facts to mean there are two different specifiers for subjects in German (as well as in Old English). Specifically, he argues that subjects (all subjects, in OV Germanic) can appear in either Spec(AgrSP) or Spec(TP). However, the evidence Haeberli considered really does not support a conclusion more

fine-grained than that there are two positions for subjects, one higher in the structure and one lower; the precise identities of the two positions are not clear. Examining the example sentences above more carefully, the fact that the lowest subject position is to the right of manner adverbs tells against Haeberli's conclusion that the low position for subjects is also a derived position, and therefore above the subject's original *v*P. Certainly the idea that the low subject position is Spec(TP) does not appear to be correct. I would suggest that the low subject position is simply the subject in situ and the high subject position is the scrambling of subjects to a TP-adjoined position. It is, of course, not a coincidence that the subjects in the examples above are indefinite; definite subjects do not remain within the *v*P in the same way (the well-known "definiteness effect"). Rather than treating this as a stipulated condition on the low subject position, as Haeberli (2005) does, the analysis here brings this fact in line with the strong general tendency for definite subjects to scramble in German (Diesing 1992, 1997).

Thus, the two subject positions argued for by Haeberli are really just an unscrambled and scrambled position for subjects. If this is correct, then the "definiteness effect" found in VO Germanic (e.g. English, or Icelandic as discussed by Rögnvaldsson 1984a), where definite subjects cannot occupy a low subject position with an expletive in the high subject position, is in fact related to the strong tendency for definite arguments to scramble in OV Germanic. From this perspective, the definiteness effect in VO Germanic stems from a reanalysis. Presumably at some point during the stage of change from OV to VO, something that was originally a constraint on the semantic and information-structural interpretation of scrambled and unscrambled definite arguments in

the OV stage of the language was reinterpreted by language-learners as a hard syntactic constraint on the in situ subject position.

The fact that subjects may remain in Spec(vP) in tensed clauses follows naturally from a theory in which the EPP causes the entire vP (at least) to move to Spec(TP), rather than just the subject alone. Without such an analysis, the satisfaction of the EPP in the "a" sentences above is somewhat mysterious, and one must assume that an empty expletive occupies Spec(TP) in these examples which is coindexed with the subject in a lower position, even though expletive subjects in German are generally overt (*es*); this is an extension of the argument that Biberauer (2003b) made for the precursor of the current analysis with regard to presentational sentences in OV Germanic.

But beyond its basic plausibility as an analysis, movement of the whole vP makes the correct prediction that not only objects precede the finite verb alongside the subject in verb-raising clauses, as in (9) above, but also that other elements such as adverbs precede the finite verb, as in (10). Other approaches which are focused on the preverbal position of the object, on the other hand, frequently ignore this fact (cf. Zwart 1993, 1997, which make crucial use of "AgrOP" to motivate object movement to preverbal position). Such approaches derive verb-raising clauses (and indeed, all OV and Tense-final orders) by positing individual motivations for the movement of the subject and object to positions preceding the tensed verb, and in doing so, they either wrongly predict that other elements (e.g. adverbials) should be stranded in a postverbal position by default, or simply ignore the issue. Verb Raising and OV are not about movement of the object, per se, but rather about the position of the entire complement structure of the verb(s). For

that reason, approaches relying on Case to derive verb-raising and/or OV are also doomed to failure; any account relying on Case as a motivation will actually predict that non-DP constituents of the vP will be postverbal in languages like Dutch and German.

On the other hand, approaches such as Hinterhölzl (2006) do not leave themselves open to this criticism. His analysis does deal with the problem of accounting for elements other than the subject and object being in preverbal positions in West Germanic, but he does so by proposing a host of largely unmotivated movements to force each element of the vP to move individually to preverbal position (including nonfinite verbs in non-verb-raising clauses). Because each constituent of the vP must move individually to a specifier in Hinterhölzl's system, there is in principal no limit to the number of functional projections in his clausal projection. Furthermore, he is left with no way to say that preverbal position is the default position for all vP-related constituents in a language like Dutch. Finally, once we begin to address the issue of verb-projection-raising clauses, in which elements can intervene between the auxiliary and lexical verb in clauses like (29) and (30) above (in some varieties), it becomes clear that "vP-evacuation" approaches like Hinterhölzl's simply do not make interesting predictions. One simple example of this is that when a variety has both "nonfinite-verb > auxiliary" and "auxiliary > verb" orders available, and verb-projection-raising, material can only intervene between the nonfinite verb and the auxiliary when they are in the "auxiliary > verb" configuration (Hoeksema 1994). Such a strict, clear, and arbitrary prohibition seems like it should be a purely mechanical result of how these structures are derived. But in a system like Hinterhölzl's, where every element moves individually from

postverbal to preverbal position, for individual reasons to individual specifiers, this generalization can only be stipulated. On the other hand, I show below that below under the account I am developing here, this generalization simply follows as a consequence of the formal system.

6.3.3 Three-verb clusters in West Germanic

One strong piece of evidence in favor of the approach taken here, in which the EPP, verb-raising, and verb/head-finality are all in some sense the same thing, is the distribution of verb-raising patterns in 3-verb clauses. In subordinate clauses containing two auxiliaries, there are 3!, or six, logically possible orders of the three verbs. As Wurmbrand (2004, 2005) states, based on her survey of the verb-raising literature and original empirical investigations, only 5 out of 6 of these orders are actually attested in the OV West Germanic varieties. Two of these orders are simply generalizations from patterns I have already discussed. The first is the standard German order that I have already discussed in the preceding sections and chapter (the "3-2-1" order, following Wurmbrand's terminology), which is derived by the pied-piping of the maximum amount of structure at each step of the subject's movement to Spec(TP); as we have seen, this results in a complete "roll-up" of the clausal structure. The second order is the generalization of the standard Dutch pattern I discussed above, which is the mirror image of the standard German order (Wurmbrand's "1-2-3" order); this pattern is derived by the

cyclic movement of the subject's *v*P through the specifiers of other functional projections until it reaches Spec(TP) with no further pied-piping, leaving the three verbs in their VO-like base order as in the two examples below.

West Flemish

(19) da Jan geen vlees hee wilen eten
that Jan no meat has (1) want (2) eat (3)
"…that Jan has not wanted to eat any meat"
(Haegeman & van Riemsdijk 1986: 442)

Early Yiddish

(20) dz ikh nit hab velin ab lasin
that I not have (1) want (2) off let (3)
"…that I didn't want to stop"
(Weinryb 1937, from a Yiddish letter written in Cracow, date: 1588; LE2,51.16 in the *Penn Yiddish Corpus*, Santorini 1997/2008)

The remaining three possible verb-raising orders are illustrated in the examples below.

1-3-2
Afrikaans

(21) dat Jan Marie kan gesien het
that Jan Marie can (1) seen (3) have (2)
"that Jan could have seen Mary."
(Robbers 1997)

German

(22) daß er hätte kommen können
that he had-SUBJ (1) come (3) can (2)
"that he would have been able to come"
(Haegeman & van Riemsdijk 1986: 427)

2-3-1
Afrikaans

(23) dat Jan kon werk het
that Jan could (2) work (3) has (1)
"that Jan has been able to work."
(Robbers 1997)

3-1-2
Swiss German
(24) dass er vorsinge wil chöne
that he sing (3) wants (1) can (2)
"that he wants to be able to sing"
(Wurmbrand 2005: 1)

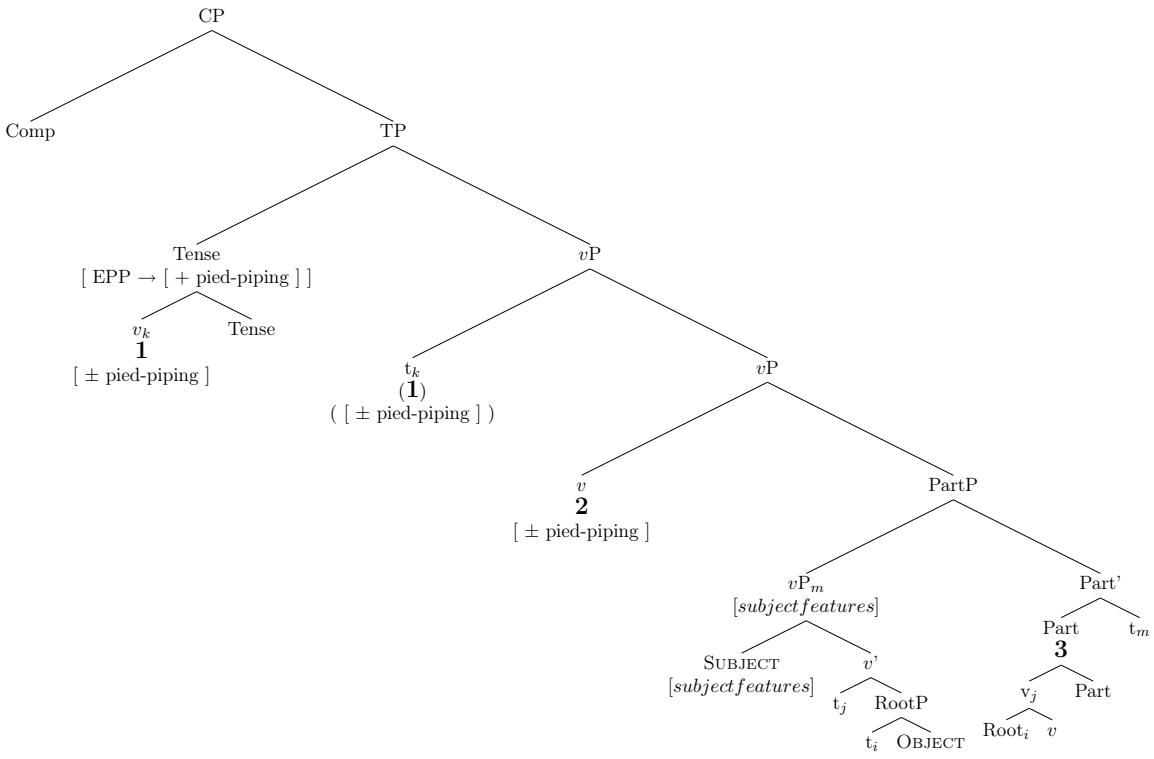
Early Yiddish
(25) ver nit loyfin hat kenin…
who not run (3) has (1) can (2)
"Who was not able to run…"
(*Der tsveyter khurbn fun Ukrayne* ("The second destruction of the Ukraine"); date: 1783; UK1,33 in the *Penn Yiddish Corpus*)

The one unattested 3-verb pattern across the verb-raising languages is the *2-1-3 order, in which the finite verb is sandwiched between nonfinite verbs (Wurmbrand 2004, 2005); this order was also consistently rejected by the individuals surveyed in Wurmbrand (2004), who were sampled from across the South German dialect continuum.[67]

The current approach to verb-raising predicts that only 5 out of 6 logically possible combinations of 3 verbs are possible across the varieties of West Germanic, and it predicts the correct 5 patterns. The first two possible orders, 3-2-1 and 1-2-3 have already been addressed. Taking as our starting point the clausal schema below, with places for two auxiliaries and a lexical verb, the other three grammatical orders are

[67] The only exception appears to be a single speaker from Switzerland who seems to accept some sentences with a 2-1-3 order some of the time. Wurmbrand (2004: 9, 12) does not consider these isolated judgments to be robust enough for her to abandon the generalization that *2-1-3 is a general fact about verb-(projection)-raising.

derived with different distributions of [± pied-piping] features on the "1" and "2" heads, which are marked "[± pied-piping]" in the tree below.[68]



The schema above shows the derivation after head-movement and after the movement of the lowest *v*P to Spec(PartP), but before the subject has moved any farther on its way to Spec(TP) and before any further feature-percolations have taken place.

The three grammatical verb-raising orders shown in (35)-(39) are derived in the following way. The 1-3-2 pattern results if the functional heads controlling pied-piping have the pied-piping features distributed in this way:

---

[68] The feature underneath the trace of *v*-to-T movement, t_k, is in parentheses only to indicate that it doesn't matter if the "1"-auxiliary's pied-piping feature is satisfied at its base position or at its derived position. There is only meant to be **one** pied-piping feature corresponding to this head.

1: [– pied-piping]
2: [+ pied-piping]

In the 1-3-2 derivation, the "2" head triggers pied-piping by percolating the subject-features on vP$_m$ one level higher to PartP. The next step of the subject's movement up the tree accordingly pied-pipes PartP to the specifier of the 2-auxiliary's vP. At this point, the [–pied-piping] feature on "1" must be satisfied, and that happens by deleting the percolated subject-features on PartP, leaving them only as high as vP$_m$, as they were at the beginning of the paragraph. In the final step of the derivation vP$_m$ is attracted to Spec(TP), resulting in the 1-3-2 order of verbs with all of the constituents of vP$_m$ preceding the finite auxiliary.

The remaining two orders both involve some underspecification for [pied-piping] on the part of one of the heads involved. The 2-3-1 order is the result of applying the following featural specification to the clausal schema:

1: [+ pied-piping]
2: []

The "2" verb is unspecified for pied-piping, which means that the vP$_m$ will move cyclically past "2" to its specifier. It is now in the specifier of "1"'s complement, in position to satisfy the 1-auxiliary's [+ pied-piping] feature by percolating vP$_m$'s subject-features to vP headed by "2". This causes the 2-auxiliary's vP to be pied-piped with the subject to Spec(TP), taking with it both vP$_m$ in "2"'s specifier and "2"'s complement, PartP, placing the 2- and 3-verbs in front of the finite auxiliary "1", in their original 2-3

order. The remaining order, 3-1-2, is the consequence of the opposite feature assignment from the last one:

1: []
2: [+ pied-piping]

The 2-auxiliary causes the subject-features to copy from $v\text{P}_{\text{m}}$ to its dominating PartP, and this is the only feature copying or deletion that's required (after $v\text{P}_{\text{m}}$ already has its copied subject-features) to produce the correct final structure for 3-1-2. Once PartP has the copied subject-features, the feature on the 2-auxiliary is satisfied. The PartP is then simply attracted cyclically up the tree to Spec(TP), correctly placing all of the constituents of $v\text{P}_{\text{m}}$ and the nonfinite verb heading PartP before the 1- and 2-auxiliaries, which remain in their original positions.

Lastly, the analysis excludes the ungrammatical *2-1-3 order without stipulation. According to the antisymmetric analysis of head-final languages, the *v*P headed by the 2-auxiliary must dominate the 3-auxiliary and its projection at the initial merge stage of the derivation. This is shown in the clausal schema above, and the reader will see that no derivation progressing from that starting point can move "2" without "3" by XP-movement, and there is no independent motivation to achieve this by head-movement. It is just not possible to pied-pipe the 2-auxiliary's *v*P to Spec(TP) without also bringing the 3-auxiliary along with it; there is no constituent containing "2" but not "3". The ungrammaticality of *2-1-3 thus follows from the mechanics of the system and the assumptions about the base order which are merely part and parcel of a unified,

antisymmetric account of Germanic. The fact that exactly 5 out of 6 logically possible orders are predicted by this analysis, and the 6th order is the only unattested one cross-linguistically, is a strong indication that this analysis is on the right track at a surprising level of detail. And again, this result simply falls out from the analysis, without the addition of any machinery to exclude it.

6.3.4 Verb Projection Raising

In addition to the verb-raising orders discussed above, the approach above also predicts the existence of the Verb Projection Raising (VPR) construction, when the [pied-piping] feature system is combined with the analysis of scrambling as constrained adjunction. Languages such as Swiss German (Haegeman & van Riemsdijk 1986, Penner 1990), West Flemish (Haegeman & van Riemsdijk 1986), Afrikaans (Robbers 1997), Old English (van Kemenade 1987, Pintzuk 1991, Haeberli & Haegeman 1998, Pintzuk & Haeberli 2006), and Early Yiddish (Santorini 1992, 1993a) display the kind of variation shown below in clauses that are clearly Tense/Infl-final (in classical terms) under some independent criterion.

West Flemish:

(26) a. da Jan een hus kopen wilt
that Jan a house buy wants
"…that Jan wants to buy a house"
b. da Jan een hus wilt kopen
c. da Jan wilt een hus kopen

Zurich Swiss German

(27) a. das de Hans es huus chaufe wil
that Hans a house buy wants
b. das de Hans es huus wil chaufe
c. das de Hans wil es huus chaufe
(Haegeman & van Riemsdijk 1986: 419)

The first two cases above, "a" and "b" sentences, have already been dealt with in this dissertation: the first case is the German case, in which the maximum amount of structure has been pied-piped, and the second is the standard Dutch case, in which only the *v*P immediately dominating the subject has been pied-piped to Spec(TP). The "c" sentences are the VPR case, in which a constituent of the lower vP appears between an auxiliary and its governing lexical verb. Another example from Early Yiddish appears below. Note that the negation, *nist*, preceding the finite auxiliary means that this clause is Tense-final in the way that all Dutch and German clauses are (see also example 34 above). Under the antisymmetric analysis, this means that the vP has moved to Spec(TP), taking with it the diagnostic negation.

(28) dz er nist hat kenn **fun mir** brengn
that he not has could from me get
"…that he has not been able to get [it] from me"
(*Court Testimony*, West Yiddish, date: 1565; ID CO,68 in Santorini 1997/2008)

I propose that the third, "c"-type of case is like the second, "b" case, but with the option of scrambling exercised at an intermediate point in the derivation. Rather than scrambling occurring only after the subject has landed in Spec(TP), pied-piping its *v*P,

the formalism I have proposed makes it possible for scrambling to occur as the *v*P moves cyclically up the clausal spine. The Conservation of C-Command (i.e. the GHC) allows scrambling in principal at any point of the derivation, provided that it is respected in the configuration which has been produced at the point when the scrambling occurs. In other words, intermediate scrambling can occur as long as the vP has escaped the c-command domain of the head which projects the adjunction site of the scrambling element. Some of the possible surface strings resulting from intermediate scrambling are shown in the paradigm from Bernese German below (verbs are in boldface):

(29) a. … dass er **het wöue** sire Frou es Gschänk **mache**
that he **has want** his wife-DAT a present-ACC **make**
"…that he wanted to give his wife a present

b. … dass er **het** sire Frou **wöue** es Gschänk **mache**
c. … dass er **het** sire Frou es Gschänk **wöue mache**
d. … dass er sire Frou es Gschänk **het wöue mache**
e. … dass er sire Frou **het** es Gschänk **wöue mache**
(Penner 1990: 168)

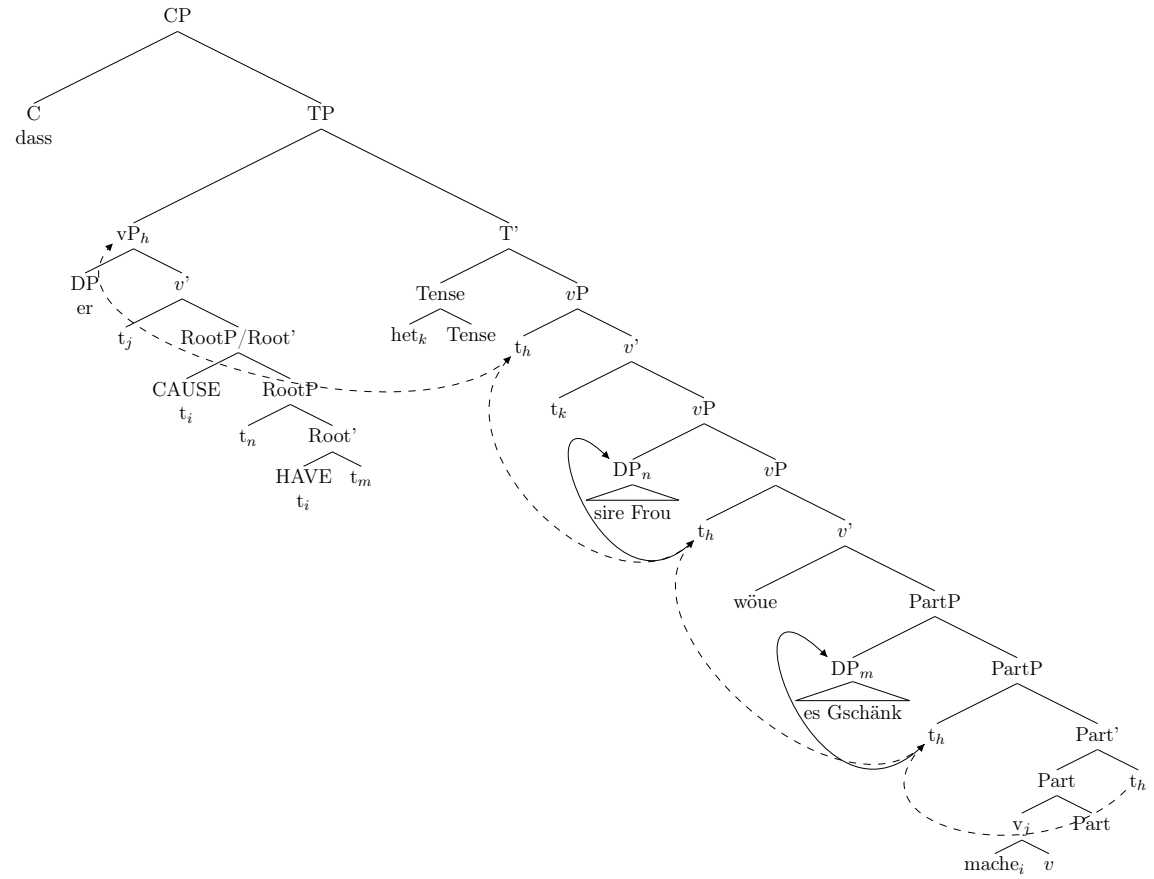
The analysis here derives all of the orders in (36) from independent principles of the system. Bernese German differs from other West Germanic varieties in that V(P)R is obligatory with all auxiliaries and modals (in fact, all verbs which take a complement that is in the extended verbal projection) except for perfect HAVE and BE, with which it is optional (Penner 1990: 168). The generalization is very easy to state in the present framework. Abstract Tense is a [+ pied-piping] Probe across the board in Bernese German, most flavors of little-*v* head (i.e. auxiliaries) are unspecified for [pied-piping], and perfect *ha* ("have") and *si* ("be") show two variants: a [+ pied-piping] variant and a

variant that is unspecified, or inert, with respect to pied-piping. When all of the heads on the clausal spine (except for Tense) are inert with respect to pied-piping, the *v*P containing the subject moves cyclically from specifier to specifier as in (32) above. The only difference between the Dutch structure in (32) and the Bernese German sentences in (36) is that in (36a), (36b), (36c), and (36d), the direct object, indirect object, or both scramble out of the subject's *v*P at some point on its path up the clausal spine. In the tree below I show the structure for (36b) below in order to illustrate the analysis.[69,70]

[69] The tree here is not meant to indicate a stance on the question of whether the bare infinitive-taking verb *wöue* ("want") takes a *v*P complement or a TP complement. I have chosen to represent its complement as a *v*P for ease of explication, but the details do not change the basic proposal.

[70] The ditransitive structure in the tree below follows Harley (2002) in general, though nothing in the present discussion depends on these details. The "CAUSE" and "HAVE" labels are only meant to indicate where the decomposed heads making up the ditransitive structure originally were; of course, these heads have already moved and incorporated with a number of functional heads at this point in the derivation, as indicated by the traces.

(30)



The solid movement arrows in the structure above show at what point in the derivation the two objects have scrambled out of the *v*P as it moves cyclically up the tree. The direct object, *es Gschänk*, moved to adjoin to PartP when the subject's *v*P was in Spec(PartP), and the indirect object *sire Frou* scrambled to adjoin to the *v*P headed by *wöue* when the subject's *v*P was in the specifier of that projection.

The current analysis of VPR also explains a well-known asymmetry in the West Germanic verb-cluster data: in languages which allow some "head-final" orderings among verbs and some verb-raising orderings, it is only the verb-raising, "head-initial" or

VO-like orderings that allow constituents to be placed between the various verbs (verb **projection** raising); even in the same language or dialect, the standard German-like "head-final" orderings do not allow constituents to intervene between the verbs (Haegeman & van Riemsdijk 1986, Hoeksema 1994, Wurmbrand 2005). This is the type of asymmetry between "head-final" and "head-initial" orders that one would expect to see if Kayne's antisymmetry hypothesis is correct. In fact, no symmetric view of phrase structure would predict such an effect *a priori*; it would predict the opposite, in fact. Similarly, any account which relies on head-movement or the morphological reorderings of heads by definition cannot handle any correlation between the surface order of heads and projection-raising possibilities (Wurmbrand 2005: 24) also makes this point); such an account is also too "symmetric", in a sense, regardless of which basic order it assumes, because it assumes the various orders are hierarchically identical. As it turns out, the antisymmetric, XP-movement account here captures the generalization as a result of the mechanical system we've built up to this point in the discussion, without any additional stipulations.

When the verbs are in the V(P)R order, it means that some subconstituent of the clause (at least the subject's *v*P) has moved cyclically past them, which leaves the possibility open for something to scramble out and adjoin between the verbs. But anytime that the verbs are in the "head-final" non-V(P)R order, the leftmost verb only appears in that position by virtue of its entire projection (and anything it dominates) being moved leftward to the specifier of the rightmost verb. Anything adjoined to the XP which moves leftward would also have to move leftward because the additional XP node

created by the adjunction would be identical to the original XP node, including any copied features (which trigger the leftward pied-piping of the XP). Note that the generalization holds true whether all of the verbs in the clause are in the verb-raising order, or only some of them are. For instance, the analysis here predicts that "raised"-constituents should be able to intervene in VPR clauses with the 2-3-1 order, but only between the "2" and "3" verbs, not between the "3" and "1" verbs. This prediction is correct (Wurmbrand 2005), and is illustrated in the example below:

(31) da Valère willen Marie dienen boek geven eet
that Valère want (2) Mary this book given (3) has (1)
"that Valère has wanted to give Mary that book."
(Haegeman 1998: 260)

Similarly, as predicted, the 1-3-2 order allows material to intervene between the "1" and the other verbs, but not between the "3" and "2". This is shown by contrasts in the standard German and Swiss German examples below:

German

(32) a. daß er das Buch hätte genau durchsehen sollen
that he the book had (1) carefully through-look (3) shall (2)
"that he should have looked through the book carefully"

b. * daß er das Buch hätte durchsehen genau sollen
that he the book had (1) through-look (3) carefully shall (2)
(Zwart 1996b)

Swiss German

(33) a. ob si hett d Prüeffig besto chöne
whether she had (1) the exam pass (3) can (2)
"[who knows] whether she would have been able to pass the exam"

b. * ob si hett besto d Prüeffig chöne
whether she had (1) pass (3) the exam can (2)
(Wurmbrand 2005: 23, citing Haeberli p.c.)

Crucially, each scrambling movement obeys the Conservation of C-Command constraint when it takes place, according to the c-command relations holding between the scrambling object and functional heads at that particular point in the derivation. The role of the Conservation of C-Command in the derivation of VPR is dramatically illustrated by an outstanding natural example found by Haeseryn (1990: 81):

(34) W. wurgde zijn echtgenote, nadat hij er was **achter** gekommen
W. strangled his wife after he there was **behind** come

dat ze een verhouding had.
that she an affair had
"W. strangled his wife after he had found out she was having an affair"
(*De Volkskran*, 10/5/1988)

Hoeksema (1994: 5) cites (and translates) the example above from Haeseryn, and points out that the postpositional PP in the subordinate clause *er achter* (lit. "there behind") is surprisingly split by VPR. Only the P head *achter* (in boldface above) appears between the two verbs in the subordinate clause in the VPR, while its object appears in the canonical preverbal position. This type of VPR pattern is entirely expected under this account, just in case the PP is postpositional. In that case, the complement of P has presumably moved to Spec(PP), placing the object *er* out of the (asymmetric) c-command domain of P. Assuming this is correct, the object *er* is free to scramble and escape the PP

because no violation of the Conservation of C-Command will ensue if the object is scrambling from Spec(PP). Suppose the object has done so, and adjoins to the *v*P containing the subject. The subject's *v*P is pied-piped up the clausal spine in the usual cyclic way for V(P)R, but when it reaches Spec(PartP), the now remnant PP scrambles out of the *v*P to adjoin to PartP. As I've already discussed, this movement also obeys the Conservation of C-Command because the scrambling is quite local and does not change the c-command relations which hold between the heads on the extended verbal projection and the PP at this stage of the derivation. The subject's *v*P then simply continues its cyclic movement until it reaches Spec(TP), stranding the remnant PP, but bringing with it the PP's object which adjoined to the *v*P earlier.

It is likewise crucial for this account of VPR that scrambling is a type of adjunction. It is the adjunction that provides a landing site at which elements may be stranded between the verbs in the VPR construction. It would not be possible to maintain an account of scrambling which relies on there being a large number of dedicated specifiers of other (invisible) functional projections for elements to scramble to (as in Hinterhölzl 2006, *inter alia*), and simultaneously keep an analysis which straightforwardly relates scrambling to VPR and VPR to the non-verb-raising (standard German) orders. In order for individual heads to control how much structure below them is pied-piped as the subject moves to Spec(TP), there must be a local relationship between the given head and the subject's features; the features can only be percolated one level up by a given head, and so they can only be one projection away from the head. If there were invisible specifiers along the clausal spine, that would interfere with this

locality relationship in cases where a maximum full roll-up of structure must occur. One could salvage such an analysis by loosening the relationship between the individual heads and the pied-piping system, and instead propose to control the pied-piping with some higher level, global mechanism.

But any analysis where pied-piping is controlled non-locally would fail to capture the generalization that the distribution verb-raising and VPR constructions are well-known to be sensitive to the specific auxiliary verb(s) and morphological form of the lexical verb which occur in the clause (see Wurmbrand 2004, 2005 for an overview, but this observation has been repeatedly made in the literature going back to the original study of Dutch in Evers 1975). In other words, the specific types of functional verbal heads like little-*v* and Part (including subtypes of e.g. little-*v*) which make up the clausal structure definitely matter to V(P)R, since it is these that are spelled-out as the various auxiliaries and pieces of verbal morphology (under many current frameworks, esp. Distributed Morphology). As I remarked in the previous chapter with reference to McCloskey (2000), the possibility that different heads are specified for different values of [pied-piping] makes the prediction that the choice of functional head will have an effect on how much structure is pied-piped, and therefore the combination of different types of heads will result in different relative orderings of heads and their complements. This is likely the correct prediction for every variety of West Germanic which shows V(P)R orders, but it is only clearly demonstrated in those varieties which show most of the different possible VPR orders depending on the choice of auxiliaries and lexical verbs.

An extreme case is West Flemish, as Wurmbrand (2004: 5-6) points out, which restricts each of the 5 possible orders of three-verb clauses (see discussion of three-verb clusters above) to a specific combination of auxiliaries and lexical verb forms. This type of specialization only makes sense if the various V(P)R surface orders emerge from the combination of certain functional heads with their own individual feature specifications. Just as an example of how the analysis here makes the correct predictions in a case like West Flemish, take the two West Flemish constructions involving participles. Using the "1, 2, 3" notation from Wurmbrand (2004, 2005) and above, the two constructions are: "Mod-Aux-V" where 1 = a finite modal, 2 = a non-modal bare infinitive auxiliary, and 3 = a participle lexical verb, and "Aux-Aux-V" where 1 = a finite non-modal auxiliary, 2 = a non-modal participle auxiliary, and 3 = a participle lexical verb. "Mod-Aux-V" can occur in the orders 1-3-2 and 3-1-2, and "Aux-Aux-V" can occur in the orders 1-3-2 and 3-2-1.

What do these two constructions have in common? According to the distribution of [± pied-piping] features I discussed above, all four orders involve a "2"-auxiliary which is specified [+ pied-piping]. This is easy to characterize in the grammar of West Flemish by stating that all auxiliaries which select for a PartP headed by a lexical verb (i.e., a Part+Root complex head) are [+ pied-piping]. Now, what is the difference between the two constructions? They both share the 1-3-2 order, but Mod-Aux-V also allows the 3-1-2 order, while Aux-Aux-V does not allow that order but rather allows 3-2-1 instead. In terms of the theory of pied-piping features, this means that the two constructions differ in the feature specification of "1": in Mod-Aux-V, "1" is unspecified

for [pied-piping] (i.e. inert, []), and in Aux-Aux-V, "1" is either [+ pied-piping] or [– pied-piping] but never inert. How can we account for this difference in the feature content of "1"? The obvious answer is that "1" is a modal in Mod-Aux-V, and so the generalization must be that modals are always inert with respect to pied-piping (specified: []) in West Flemish. This prediction is borne out by all of the constructions involving modals in West Flemish (cf. Table 1 in Wurmbrand 2004: 5): they only allow 1-2-3 when the modal is "1", and 1-2-3 or 2-3-1 when the modal is "2", both of which involve an inert 2-auxiliary in the analysis presented here.

Finally, I will note that the analysis of verb-raising and projection-raising in particular that I have presented makes a prediction which appears to be, on the face of it, overly permissive: since projection-raising is has been reduced to the combination of cyclic *v*P movement and scrambling, there is no clear way to generate a language which has verb-raising but not projection-raising, as is sometimes claimed for standard Dutch. But this is not actually a drawback. This aspect of the system leads to the strong hypothesis that verb-raising can **never** exist in a grammatical system without VPR, and it is precisely this prediction that Hoeksema (1994) argues for. First, Hoeksema points out that people who are supposedly speakers of a standard, non-projection-raising variety occasionally produce projection-raising clauses which other purportedly standard speakers find to be acceptable. One such case is the example I have already discussed above in (48), which also happened to come from a newspaper article, a rather standard stylistic register. The example below from Den Besten & Broekhuis (1992: 25) also

shows that projection-raising is possible in standard German in rare contexts where the verb-raising order can occur:

(35) dass er noch muss **nach Bonn** zurückfahren können
that he yet must **to** **Bonn** back-drive can
"that he must still be able to drive back to Bonn"

Note that this clause follows the VPR pattern we established above for languages like Swiss German, West Flemish, etc., in two important respects: first, the position of the adverb *noch* ("yet") before the tensed auxiliary *muss* shows that the smallest *v*P has fronted to Spec(TP), providing a preverbal adjunction site for the adverb. Secondly, the constituent intervening in the verb cluster can only occur between verbs which are in the "head-initial" order relative to each other (between the "1" and "3" in this "1-3-2" clause). As predicted by the analysis here, *nach Bonn* cannot occur between the "3" and "2" verbs because the roll-up of structure that derives the 3-2 order does not leave any adjunction site in that position:

(36) * dass er noch muss zurückfahren **nach Bonn** können
that he yet must back-drive **to** **Bonn** can
(Den Besten & Broekhuis 1992: 25)

Interestingly, the main argument Hoeksema (1994) presents for the hypothesis that verb-raising implies VPR comes from his diachronic study of Dutch. First, Hoeksema's (1994: 24-25) quantitative study of VPR over the history of Dutch shows that the frequency of projection-raising remained essentially stable from the 13th through

17th centuries, but then dropped precipitously in the 18th and 19th century texts. While this is not clearly evidence bearing on the nature of the change, it is an indication that the course of the change may have been unusually erratic and should be investigated further to determine if the change was caused by a new prescription in written Dutch or some other extragrammatical effect. Secondly, Hoeksema shows that in the period after the frequency of projection-raising dropped in the 18th century, the examples which remained were of a very restricted set of types involving very short constituents. For instance, while there were examples of PPs intervening between verbs, PPs containing DPs consisting only of bare nouns were considerably favored over PPs containing DPs with overt determiners (the former were 97/132 = 73.5% of the examples containing PPs).[71] Hoeksema takes this to reflect a prescriptive pressure against the interruption of verb-clusters which continued to restrict the size of the intervening constituent as northern, written Dutch became more standardized over time (the southern varieties do not reflect this standardization in the same way). If his analysis is correct that the diachronic effect is one of language use in a context of standardization rather than one of true grammatical change, then modern standard Dutch is not really a non-projection-raising language, and the remaining restricted cases of interrupted verb clusters in modern Dutch are actually representative of a grammatical option which would be more widely used if a strong stigma or stylistic effect were not present.

---

[71] Hoeksema unfortunately does not report the number of non-projection-raising clauses containing PPs of each of these types as a control. However, if we assume that the frequency of these two types of PPs have no *a priori* reason to be different overall in Dutch and English, we have some reason to believe that the effect Hoeksema reports is real: in finite subordinate clauses with auxiliaries in the *Penn Corpus of Early Modern English*, PPs containing a DP without a determiner occur at 52.4% (1952/3719) and PPs containing a DP with a determiner occur at 47.6% (1767/3719).

6.3.5 A note on the West Germanic variation with respect to V(P)R

Before ending the discussion of VPR, I would like to make a final note about the interspeaker and intraspeaker variability of VPR. While the variation is considerable, it is not without limits. The analysis presented here predicts that this kind of variation should exist, and at the same time, it also predict that certain types of variation should not exist in Tense-final (in the relevant sense) West Germanic. The fact that there is a great deal of variation is evident in the subtle differences reported in the literature for different dialect areas. For example, it's clear that "Swiss German" is nothing like a uniform syntactic dialect, as there are differences reported in the possible VPR patterns for Zurich German (Haegeman & van Riemsdijk 1986), Bernese German (Penner 1990), and St. Galler German (Schönenberger 2001), among others. Additionally, the questionnaire study of geographical dialects in Wurmbrand (2004) and the experimental study in Bader, Schmid, & Häussler (2009) establish two additional results: first, that the interspeaker variation is much more widespread than any traditional division of geographical dialects suggests. and secondly, that individual speakers frequently (maybe usually) accept more than one V(P)R order for any given collection of verbs in a subordinate clause. This is especially true for sequences of verbs ≥ 3, and Bader, Schmid, & Häussler (2009) makes it especially clear that even self-reported speakers of "standard German" actually accept a much larger number of possible orders than they would normally produce in literary or

very self-conscious social contexts. Their study supports Hoeksema's (1994) suggestion that much of the rigidity in VPR reported for "standard Dutch" and "standard German" in the literature is actually artificially produced by strong prescriptive pressures on speakers in certain geographical areas and/or socioeconomic contexts.

The approach here predicts the observed type of variation, but not every conceivable type of variation. Since the all of the patterns found in West Germanic verb clusters emerge from the interaction between the local [pied-piping] features on the various verbal heads, any variation in these patterns is just variation in the featural content of some head or heads. In other words, it is morphosyntactic variation that is no different from speakers differing in whether a given head shows agreement or not; indeed, exactly that variation is found in the complementizer systems of continental West Germanic. Similarly, since it is possible for speakers' e-language to show morphological doublets, e.g. the English variation in *dragged~drug* for the past tense of the verb DRAG, it should also be possible for speakers to learn more than one version of a given verbal head with respect to pied-piping, e.g. a [+pied-piping] passive little-*v* and a [–pied-piping] passive little-*v*. In this way, the variation in VPR supports the hypothesis of Kroch (1994) that all syntactic variation can be reduced to variation in the feature specifications of morphosyntactic heads.

However, we have already seen in the discussion of 3-verb clusters that not every logically possible type of variation is attested: the *2-1-3 order is not a possible variant, and the analysis here correctly rules it out. In a similar way, while my analysis predicts a potentially large amount of variation in the types of constituents that can intervene inside

the verb cluster in VPR, depending on which elements scramble out of the *v*P at an intermediate step in the derivation, the "stranding" of XPs is also not unconstrained. As we have already seen, XPs can only intervene between heads in the "head-initial", verb-raising order, not those in the "head-final" order, and this holds true for whatever set of orders a given speaker has in their grammar. It is also true that constituents cannot be freely "stranded" to the right of the whole verb cluster, with the exception of a few restricted extraposition constructions (see the chapter below on Heavy NP Shift for a treatment of one of these). Note that an analysis like the one in Hinterhölzl (2006) can also extend to cover the observed variation, but it has no clear mechanism for restricting the possibilities for variation to the ones which are actually attested. In a "VP-evacuation" system like Hinterhölzl's, a head-final structure is derived from a head-initial base by the movement of all *v*P-related elements individually to pre-Tense positions, including nonfinite verbs. There are a large number of specifiers preceding Tense in the structure, and each of these may be filled by some *v*P-related element, and some by more than one different *v*P element, producing the various surface orders. In such an approach, ruling out any particular order, e.g. *2-1-3, becomes a matter of ad hoc stipulation with some diacritic. The projection-raising orders are derived by stranding some XP in its base position to the right of the finite verb. But if XPs can be freely stranded in their base position, then there is no theory of why some elements are **never** stranded in OV West Germanic, e.g. light adverbs. Indeed, adverbs should be the most frequently stranded item in Hinterhölzl's account, since their presence implies another level of structure (either an AdvP in the clausal spine or by adjunction to *v*P, etc.); since adverbs add a

level of structure, the original *v*P that the adverb is modifying should be able to front to pre-Tense position by itself to satisfy whatever feature it needs to, stranding the adverb at the end of the clause. This is not possible in the analysis here, since the entire structure to the right of the final verb is pied-piped in OV West Germanic via feature-percolation; there is no structure left after the last verb to attach an adverb to. In contrast to a VP-evacuation-style proposal, the restrictions on the possible variation I have just noted follow immediately from the formalism I have developed in this dissertation thus far.

6.4 Summary and Conclusions

In this chapter I have shown that the pied-piping approach to head-finality, coupled with scrambling-as-adjunction and the Conservation of C-Command, yield a clean account of the scrambling patterns I originally discussed under the Generalized Holmberg Constraint. Recasting the constraint on scrambling in terms of c-command forced an antisymmetric account of head-final clausal structures which is not stipulative, being motivated by the EPP and resultant subject-movement to Spec(TP). More importantly, this step in turn forced us into a particular analysis of verb-raising, verb projection raising, and their relationship to the non-VPR patterns in OV West Germanic and VO Germanic which was better than previous accounts and did not require any further stipulations.

These patterns are not added to the theory; they are simply expected to exist under the theory. The predictions of the pied-piping analysis extend naturally to the case of three-verb clauses in VPR languages, and correctly predict the non-existence of the *2-1-3 pattern without additional stipulation. The fact that VPR is sensitive to the particular identity of the verbal heads in the clause makes it especially plausible that the individual heads cause the different VPR patterns to emerge by a combination of their individual featural content. As it turns out, controlling the pied-piping with local features not only fits with that observation, but it also makes the correct empirical predictions, and it does so without relying on theoretical assumptions that are any more exotic than the head-spec configuration.

In the next chapter, I explore another class of phenomena which receive a new analysis based on the antisymmetric view of phrase structure I have developed thus far. As in the preceding chapter, the main point will be the following: this new, antisymmetric way of looking at things forces us to abandon some types of analysis and adopt others, and the analyses we are forced to adopt turn out to have considerable benefits over the others.

Chapter 7

Further Consequences and Predictions: Heavy NP Shift

7.1 Introduction: why heavy NP shift is relevant to this dissertation.

Consider the now famous (with minor variations in the literature) modern English sentence in (1) below.

(1) I met on the street my rich uncle from Detroit.[72]

Going back to Ross (1967), "Heavy NP Shift" (HNPS) examples like the one in (1) have been frequently analyzed as involving rightward movement of the DP/NP (e.g. "my rich uncle from Detroit") to right-adjoin to either vP/VP or TP/IP, resulting in the order where the object follows the adjoined modifier. This section addresses the simple fact that such rightward movement accounts are incompatible with the antisymmetric view of phrase structure that this dissertation argues for. Instead, I provide an alternative type of analysis which is not only compatible with the antisymmetric approach, but also leads to

[72] Or, "Grosse Pointe."

a more productive line of work on both the syntax and information structure of heavy NP shift.

There are two main candidates for the type of analysis that must be applied to HNPS under an antisymmetric phrase structure. The first potential analysis of HNPS without rightward movement is the one outlined in Kayne (1994) itself, in which the "shifted" object is not actually moved at all; rather, it is "stranded" by the movement of some other constituent (this is the style of analysis that Kayne suggests for all of the constructions that are classically derived with some type of rightward movement). The second potential type of analysis is the one outlined (and argued against, erroneously as I'll show) in Williams (2003): the "shifted" object is fronted to some specifier position, and then the entire (remnant) clause is fronted to a specifier position above the landing site of the object. In this section, I'll argue that the latter style of analysis is the only one that can fit the data, once Heavy NP Shift is taken in the light of West Germanic as a whole. Specifically, I show that HNPS is a type of "topicalization" or focus fronting, in which the clause is separated into a focus and an open proposition (see Prince 1999 for discussion of the two focus-related functions of English topicalization); the "shifted", or postposed, object moves to exactly the same position as "topicalized" focused objects in English. The difference between HNPS and usual focus movement is that while the object is the Focus in both cases, in the latter the TP or vP is the Topic in a highly constrained version of the "Split-C" system of Roberts (2005).

The classical analysis of HNPS poses considerable empirical and theoretical problems. The most obvious theoretical problem is that a system with both left- and

right-adjunction is straightforwardly incompatible with the antisymmetric view of phrase structure in Kayne (1994) or Chomsky (1995).[73] But more relevant to the present discussion, once I have restated the GHC in purely hierarchical terms as the Conservation of C-Command, it becomes impossible to keep the Conservation of C-Command from predicting that HNPS shouldn't exist. If HNPS is an optional movement of an object to adjoin to vP or TP, then it is identical to scrambling hierarchically and is only distinguishable in terms of the (rightward) direction of the movement and adjunction. In addition to these concerns, the idea that two syntactic operations are distinguished solely on the basis of their directionality rankles with the growing consensus in the literature (since at least Chomsky 1993, if not Chomsky 1986) that the narrow syntax operates on hierarchical structure and linearization is a product of the syntax-PF interface. In fact, this is precisely the line taken in Saito & Fukui (1998), who explicitly claim that scrambling in Japanese and HNPS in English are the same operation (modulo language-specific linearization conditions, which apply at PF). However, if the GHC/Conservation of C-Command is a valid empirical generalization, then it argues against any analysis which treats scrambling and HNPS as the same thing; c-commanding heads clearly do not block HNPS under the classical rightward-movement accounts.

Finally, the classical view of HNPS as hierarchically nondistinct from scrambling suffers from (at least) three empirical problems: first, since scrambling is potentially unbounded (modulo the Conservation of C-Command), as shown by Comp-final

[73] Actually, it's worth noting that the formal syntactic architecture in Kayne (1994) does not in itself deny the possibility of rightward movement or right-adjunction. It only states that neither rightward movement or right-adjunction can exist in a language that also has leftward movement and left-adjunction. In principle, a language with only rightward movement and right-adjunction (and right-specifiers) could exist according to Kayne (1994); he treats it as an empirical fact that such a natural language does not exist.

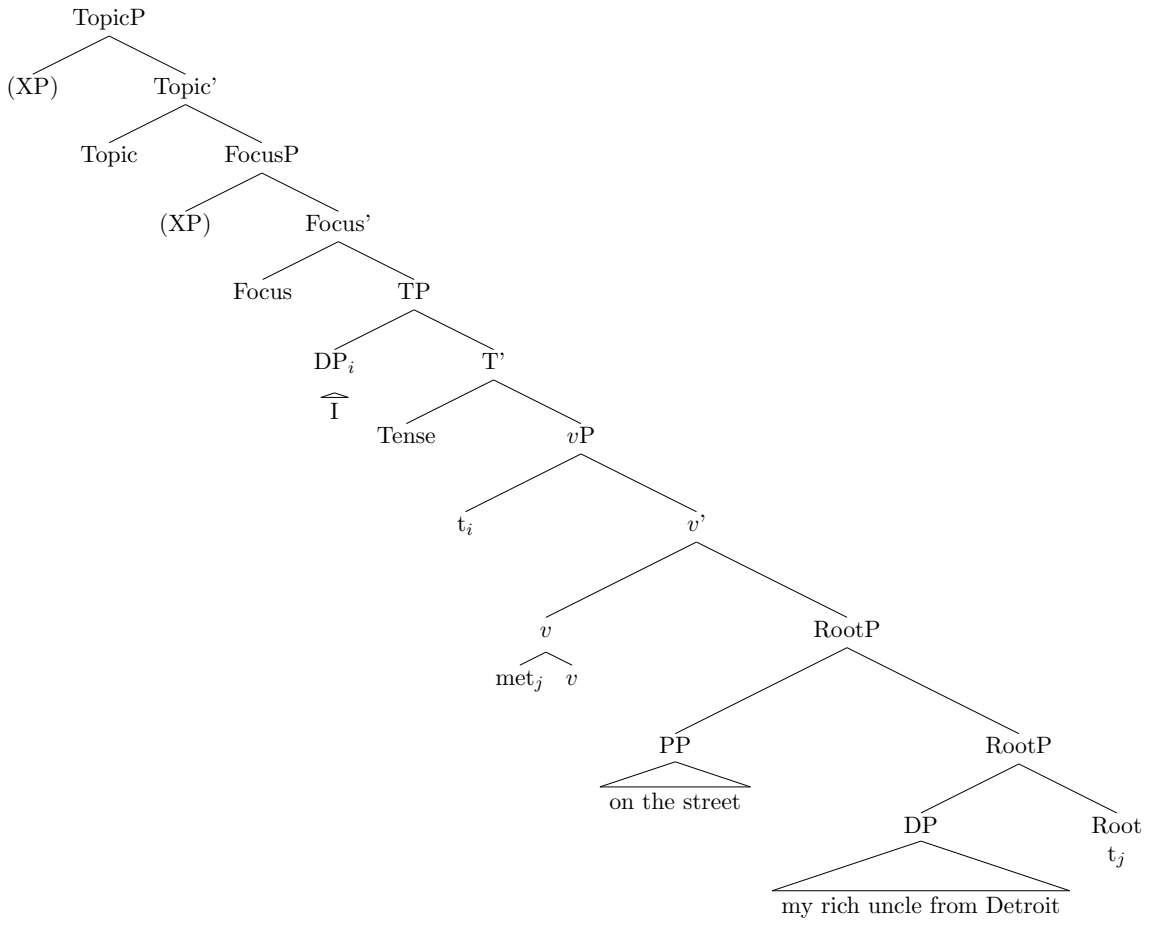
languages like Japanese and Korean, there is no explanation for the Right Roof Constraint on HNPS (Ross 1967, Baltin 1982, Baltin 2006 and references therein) in languages like English. And secondly, since scrambling to the same hierarchical position that is assumed for HNPS, vP- or TP-adjoined, cannot license parasitic gaps (as I discuss below in Chapter 8),[74] there is no natural explanation for why HNPS can license parasitic gaps in English, as in the following sentence.

(2) I met t_i on the street, without recognizing <pg> immediately, my rich uncle from Detroit.

And lastly, the third empirical problem directly confronts the strongest statement of the idea that scrambling and HNPS are hierarchically identical, the one put forward by Saito & Fukui (1998). They argue that scrambling and HNPS are identical, subject to language-specific PF linearization constraints, and so left-headed languages allow HNPS and right-headed languages have scrambling. This very clear hypothesis is falsified by the most serious empirical objection to the rightward movement accounts: as I will discuss at greater length in a moment, there exist languages with leftward scrambling *as well as* HNPS (in addition to other extraposition operations), including the primarily right-headed languages: Early New High German (and some modern German dialects), Early Yiddish, and Old English. In short, any analysis that takes into account both scrambling and HNPS at the same time must stand up to the test of the OV varieties of Germanic.

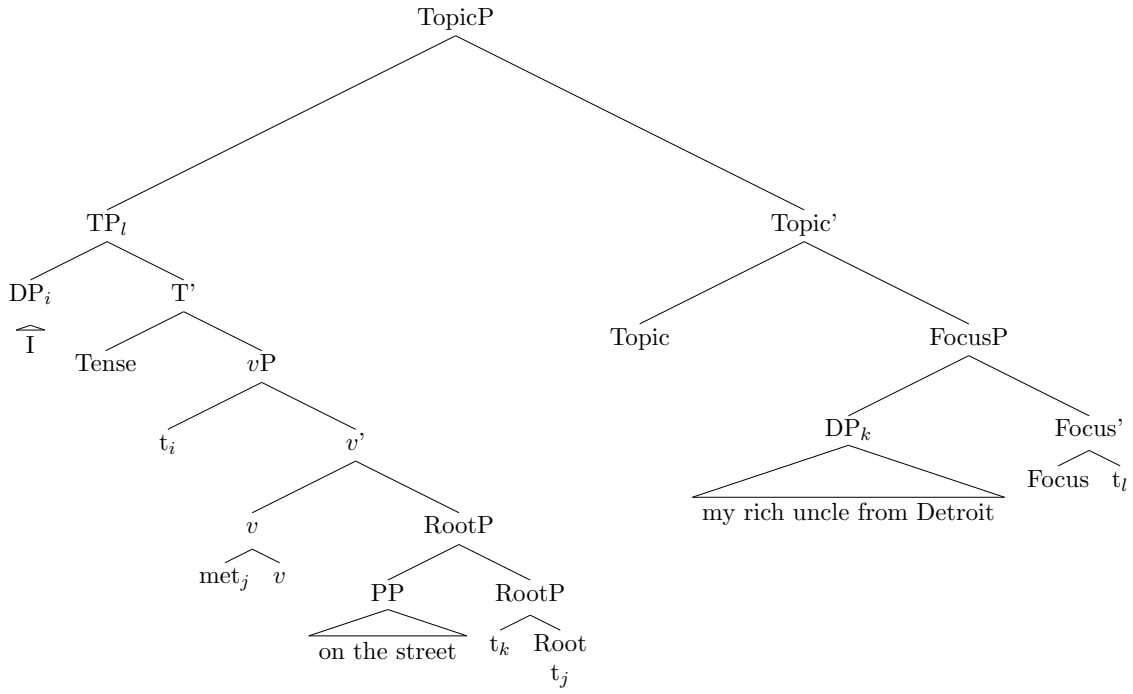
[74] As the reader will see in Chapter 7, even if this conclusion is under dispute for German and Dutch, it cannot possibly be denied for Swedish, which also happens to be the language in which parasitic gaps were discovered.

The analysis that I argue for here is that Heavy NP Shift is derived by the "NP" being fronted to a high specifier, followed by the fronting of the remnant TP to the specifier of an even higher projection, which I will ultimately show to be Spec(FocusP) and Spec(TopicP) respectively. I will explain the details of this analysis further on in this section, but the basic structure for (1) is shown below in (3). Beginning with the following structure,



we derive the order in (1) by "topicalizing" the object to Spec(FocusP) and then fronting the remnant TP to Spec(TopicP), with the final structure shown below in (3).[75]

(3)



[75] The base position (or more accurately, the initial adjunction position) of the adjunct *on the street* in these structures is an approximation and may not be correct in detail. While a full discussion of adjuncts on the right edge of the clause is beyond the scope of this dissertation, it is likely that there is more structure below the canonical position of the DP object to which the PP attaches, as originally suggested in Kayne (1994). I have not made an attempt to represent this further structure for the purposes of the two structures on either side of this footnote simply because I did not want to confuse the issue at this point in the discussion. I will note, however, that a structure in which these adjuncts at the right clausal edge are adjoined below the base position of the subject is supported by NPI licensing data such as the minimal pair below:

i. a. At no time did any student cheat on their homework.
   b. * Any student cheated on their homework at no time.

The NPI subject *Any student* is licensed by the topicalized negative PP in (ia) but it clearly is not licensed by the PP ar the right of the clause in (ib). This suggests that the adjunction position at the right of the clause is lower than even the base position of the subject in Spec(vP), which is reflected in the rough tree above.

But what is important for the discussion at hand is that the analysis of HNPS presented here will yield the same result if the PP is in a lower position, and I believe this to be a helpful direction for further research at some point in the future.

This is in sharp contrast to an analysis which assumes that the "shifted" DP remains in its original (base) position throughout the derivation and the surface order is derived by something else being moved around the DP. Any analysis along those lines cannot account for examples from OV West Germanic varieties, as shown below. These examples all show Tense-final (i.e. Infl-final) clauses which also have postposed DP objects. (Postposed objects are in boldface in the examples below.)

Early New High German

(4) es begab sich, da Jhesus vollendet hatt **solche gepott** zu seynen
it came-to-pass that Jesus completed had **such teachings** to his

tzwelff iungernn...
twelve disciples

"And it came to pass that Jesus completed such teachings to his twelve disciples"

(Martin Luther's Bible (Septembertestament), Matthew 11:1, date: 1522)

Old English

(5) Ta æfter þam þe hi gewyld hæfdon **eall heora feonda land**
Then after that-DAT that they controlled have **all their enemies' land**
"After that time when they conquered all of their enemies' land…"

(*Saint Eustace and his Companions*, date: c. 11th century, coeust,LS_8_[Eust]:388.420 in the *YCOE*, Taylor, Warner, Pintzuk & Beths 2003)

Early Yiddish

(6) in dem kll iz oykh vas an bilngn iz **zeyn hndl**
in the community is also what concern is **his trade**
"In public, [it]'s also what concerns one's business"
(*Isaac ben Eliakum's Preface to Lev Tov*, date: 1620)

Pennsylvania German

(7) Catharina Lantz Staub 1889
Catharina Lantz died 1889

[sie] Est Alt wortan **2 Eahr 8 Monat und 2 Dag**
[she] is old become **2 years 8 months and 2 days**

"Catharina Lantz died 1889. She was two years, eight months, and two days old"

(8) Benjamen Y. Lapp ist gestorben **ten 14 den May 1915**
Benjamen Y. Lapp is died **the 14th the May 1915**

[er] ist alt worden **9 Jahr 1 Mo und 12 tag**
[he] is old become **9 years 1 month and 12 days**
"Benjamen Y. Lapp died May 14th, 1915. He was nine years, one month, and twelve days old."

(Amish gravestones in Lancaster County, near Bird-in-Hand, Pennsylvania; the dates of the inscriptions are presumably the same as the dates of death)

Only the postposed object appears to the right of the verb in these examples. There are no adjuncts between the last verb and the postposed object, and there is no constituent that can be moved leftward to strand the object in final position. Each example above shows a Tense-final structure, which, under the antisymmetric approach taken in this dissertation, is derived by the movement of vP/PartP to Spec(TP) (possibly in addition to other lower movements). This does mean that part of the clause, the vP, already moves leftward past the surface position of the object, but the vP should take the object with it; there is no straightforward way to strand the object independent of the vP that contains it, and it cannot be adjoined to some invisible structure to the right of Tense, because under the most straightforward account, there is no structure remaining there in a Tense-final clause. Clearly, the analysis suggested by Kayne (1994) cannot account for this data at the same time as modern English HNPS.

7.2 Towards a unified theory of Heavy NP Shift

There are a number of reasons why we should favor a unitary account of HNPS/DP-extraposition across Germanic, in addition to the obvious methodological concerns of theoretical elegance and putting forward the strongest hypothesis possible. First, all of the Germanic varieties have had a construction like HNPS at some point in their histories, and English, Yiddish, some nonstandard varieties of German (see my discussion of Yiddish in section 2.4 above), and North Germanic (see e.g. Rögnvaldsson 1984 on Icelandic) still do have some type of HNPS in the modern languages.[76] While it is not necessarily true that these constructions are a common inheritance from historical Germanic, it is certainly a possible scenario and a hypothesis that's worth pursuing. But it becomes even more likely in light of the evidence that HNPS has not changed at all from early, Tense-final Old English to modern English. It is well-known that there is an intonational phrase boundary immediately preceding the "shifted" object in examples like (1) and (2), which also most probably indicates the clausal (TP) boundary. In Old English Tense-final clauses such as the one shown in (5) above, the clause boundary is clearly indicated by the position of the finite auxiliary. Kroch & Pintzuk (1989) showed in the Beowulf text that the position of finite verbs in Tense-final clauses which immediately precede postposed objects almost categorically correspond with half-line

[76] I will leave for further research the question of why Old High German (Robinson 1997), Middle High German (Sapp 2009), and Early New High German all clearly showed object extraposition, while modern standard (High) German does not allow DP extraposition.

boundaries in meter of the poem. Their evidence showed that there was an intonational break preceding postposed DP objects in Old English Tense-final clauses, which strongly supports the following conclusions: the prosody of Heavy NP Shift has never changed in the history of English, and the structure of HNPS (at least as far as the position of the TP boundary) has also never changed in the history of English.

Secondly, the frequency of Heavy NP Shift has never changed throughout the history of English. Pintzuk (2002), using the YCOE (Taylor, Warner, Pintzuk, & Beths 2003), presented a study of HNPS in all of the unambiguously Tense-final clauses that occur in Old English prose texts. The frequency of HNPS she calculated for Old English differs by less than 1% from the frequency of HNPS in Early Modern English, which I found from a corpus study of the PPCEME (Kroch, Santorini & Delfs 2004).[77] If we take Pintzuk's (2002) results from the entire Old English prose corpus and compare them to the study of Early Modern English, we have the results summarized in the table below.

Table 7.	HNPS	No HNPS	Total N	% HNPS
Old English Tense-final clauses (Pintzuk 2002)	123	754	877	**14.0%**
Early Modern English (dates c. 1500-1710)	730	4406	5136	**14.2%**

Essentially, the frequency of HNPS is identical at the two ends of written English history at around 0.14, and this fact is given numerical description by a chi-square statistic of

[77] The HNPS estimate for EME is based on finite clauses with an auxiliary, a nonfinite verb, a following non-pronominal object, and either a 2-word or longer AdvP or PP also following the nonfinite verb. The auxiliary condition is to exclude orders derived by V-to-T movement of finite matrix verbs and the >2-word condition excludes Germanic verbal particles which may have been tagged "Adv". In both this search and the vP-topicalization study below *Elizabeth I's Boethius* was excluded from consideration due to the unusual style and conscious archaism of that text.

0.0091, p = 0.924. In other words, under the hypothesis that time period as no effect whatsoever on the application of HNPS, you are virtually guaranteed to see a distribution of data like the one we have actually observed. Given the remarkable quantitative stability of HNPS even in the face of the multitude of phrase structure changes that took place from Old English to modern English, and no particular independent reason to suppose that the modern English and Old English constructions aren't the same one, it seems worthwhile to pursue a unitary analysis for them, and for object postposition across Germanic. By the other token, if an analysis absolutely **cannot** accommodate both Old English and modern English object postposition, such as the one in Kayne (1994), that seems to indicate the analysis is on the wrong track.

Furthermore, both modern English HNPS and the older West Germanic constructions are focus constructions, used to indicate narrow focus on the postposed object. The most notable study of the information structure of HNPS in Tense-final Germanic is Bies (1996), in which she shows that object postposition in Early New High German is very frequently used to indicate narrow focus on the "shifted" object. A more detailed discussion of narrow focus follows below, but for present purposes, a paraphrase of focus in, e.g., Vallduví (1992) would be: the specific entity in the clause that the speaker wants to draw the hearer's attention to as new information to file away in memory. Frequently, the narrow focus is an entity which contrasts with some set of entities that's already given in the discourse: it is the specific entity the speaker is talking about, as opposed to the rest of some relevant set that both the speaker and hearer are familiar with.

Another (and compatible) definition of narrow focus is: an answer by one speaker to the implied or overt question of another interlocutor. Bies (1996) uses the test of question-answer pairs for focus in order to decide if a given postposed object in an Early New High German text was in narrow focus (cf. Lenerz 1977: 20; Vallduví 1992: 64): for instance, narrow focus on the object of a given predicate, P, is an answer to the question: "What did [the Subject] P"? Such question-and-answer pairs are probably rare to hear in spontaneous speech, and certainly rare in written, literary texts. Nevertheless, one finds them occasionally, as with the example below from Early New High German, cited in Bies (1996).

(9) (Question:) Ob er auch das wort Gots predig.
Whether he [the prior] also the word God-GEN preached

(Answer:) Ja, prior hab predicirt **[F festivus diebus.]**
Yes, Prior has preached [F festivus diebus] (mass-festival of the day)
(Example cited in Bies 1996: 30)

When such explicit question-answer pairs were not available in the texts that Bies (1996) considered, she made a careful study of the surrounding context to determine which part of the clause under investigation was considered shared speaker-hearer information in the context of the preceding discourse. Bies found that where the focus of the clause could be determined from context, 46/50 or 92% of postposed objects were the narrow focus of the clause, compared with 19/142 or 13.4% of non-postposed objects (Chi-square = 98.6, $p < 2.2 \times 10^{-16}$). This result was later replicated for postposed objects in Middle High German in Sapp (2009), demonstrating that object postposition (a.k.a. HNPS)

consistently marks the narrowly focused entity in a clause during these two periods.

In fact, surprisingly, it's even possible to see the continuity between the older stages of German and a modern German dialect in the clauses taken from (relatively modern) Amish gravestones in (7) and (8) above. Gravestones represent an unusual type of discourse, in that there is a remarkable amount of "situationally evoked" information (in the sense of Prince 1981; i.e. the given entities are present in the discourse because the non-linguistic context is sufficient to make them salient in the minds of the speaker and hearer); this is of course especially necessary in this case, as the speaker (the gravestone-maker) has the double disadvantage of first, having very little he/she can say, and secondly, knowing that he/she will be potentially very far removed from the hearer/reader in time. But because there is so much situationally evoked information in a visit to a graveyard, it is quite easy to tell what the focused elements on the gravestone are. The hearer/reader knows that they're visiting dead people in a particular place and from a particular community, and that each person has a certain set of basic attributes that are normally discussed in such a context: the focused elements are the values of these attributes that the hearer/reader expects to find out from the grave marker. One such attribute is the age of the deceased: the reader- obviously knows that the given person is dead and died at a specific age, and he/she expects to find out what age that was. Therefore, the age of the deceased at death is likely to be focused, and it is postposed in the two examples above, even though Pennsylvania German is strictly a Tense-final language, just as are all the other known modern German dialects (excepting Yiddish). Similarly, the reader knows that the given person died on a certain date, but expects to be

told the specific value of the date attribute. So it is not surprising that the date of death is focused and postposed in example (8). It's also important to note that the Heavy NP Shift here is not the result of a frozen expression used in Amish gravestones: most, if not all of the other gravestones (modulo readability) in this cemetery showed the non-postposed order, whether or not they were from the same time period. The syntax here is not frozen, even if it is written in stone.

A look at some naturally occurring examples of English discourses clearly shows that the "shifted" object in the HNPS construction is frequently the narrow focus of the clause, just as Bies (1996) found for "shifted" objects in Early New High German. The clearest proof that the postposed DP is narrowly focused comes from sentences like the one below:[78]

(10) "Nothing changes tragedy into comedy like gayness. It's what we call in the entertainment world the GAY EX MACHINA."
(from the "That's Gay" feature on the TV program *infoMania*)

This example is striking because the heavy NP shift occurs within an inverted pseudo-cleft, which is well-established in the literature as a focus construction (Prince 1985, 1999, *inter alia*). It is obvious from the structure of the pseudo-cleft which element is the narrow focus, instantiating a variable in a presupposed open proposition: the variable is clearly the entity corresponding to the extracted wh-operator in the pseudo-cleft, *what*, and the open proposition is the remainder of the clause from which the wh-word was extracted. Normally, the focus in a pseudo-cleft is the other matrix clause constituent

[78] Thanks to Caitlin Light and Josef Fruewald for this example.

which is coindexed with the wh-word inside the pseudo-cleft, which in this case would be "It" at the beginning of the sentence.

However, "It" is clearly not the focus here, being an unstressed pronoun referring back to the situation described in the previous sentence; "It" is a topic, or "Ground" in Vallduví (1992). Thus, there is only one possibility left for the focus in this construction: it is the only other constituent coindexed with *what*, the heavy NP shifted *the GAY EX MACHINA*. That *the GAY EX MACHINA* is the narrow focus here is further demonstrated by the fact that it is the second object of the verb *call*, which generally renames some entity which is already salient in discourse. *Call* takes some entity which is generally known to the hearer, and provides new information about it, i.e. some new name for it. In this way, the second object of *call* is naturally going to be narrowly focused in most contexts; e.g., as in the following mini-discourse: *John bought a dog. He named it* ***Fido*** (cf. Prince 1985: 68 with regard to the verb *name*).

Additionally, as predicted by the definition of narrow focus in É. Kiss (2002), the postposed object in the examples below usually contrasts with some set of entities that has previously been evoked in the linguistic discourse, or is situationally evoked.

(11) Speaker 1: "What is this? [pointing to a food]"
Speaker 2: "It's got in it mint leaves, peanuts, and a spicy lime sauce."

In the statement made by the second speaker above, there is clearly narrow focus on *mint leaves, peanuts, and a spicy lime sauce*, and the first speaker's question could easily have read "What's in this?", the most simple question with the answer *mint leaves, peanuts,*

and a spicy lime sauce. The example below is another naturally occurring question-answer pair with HNPS, in which the speaker asks a rhetorical question and answers it with a HNPS expression.

(12) "So why are we working with TAG?
…[Because] it influences directly the kind of computations that can be used"
(Michael Collins, "TAG-based Structured Prediction Models for Parsing and Machine Translation"; talk delivered at Penn on 9/29/2009)

The HNPS clause is intended to answer the most simple paired question, "What does working with TAG change/influence?", though as we will see later, this type of example does provide a little more information than that as well.

Another example of HNPS from a natural discourse follows, this time without answering an explicit question.

(13) Context: conversation about touring lighthouses
Speaker 1: "We toured the St. Joseph one."
Speaker 2: "We toured last year…the one in Ludington."

In the example above *the one in Ludington* is new information, but it is also contrastive with not only the whole set of possible lighthouses, but also with the previously evoked entity *the St. Joseph one*. Similarly, the Heavy NP Shifted-DP in the next example contrasts with an entity linguistically evoked in the one of the immediately preceding utterances:

(14) I have two types of mosquito lotion…But I found that if you put in your pockets dryer sheets, …it keeps them away.

The set of bug-repellants is salient in the context and the object *dryer sheets*, to the right of the PP *in your pockets*, is the answer to the implied question (implied after the "but"), "What kind of repellant should I use?" Similarly, in the sense that a subset-superset relationship is involved, the example in (15) below narrowly focuses the postposed DP by identifying a subset of a set that is salient in the previous discourse.[79]

(15) Preceding context: a discussion about people who are multilingual.
Speaker: "We interviewed for a position an Australian linguist who was a perfect example of multilingualism…"

The preceding small dataset cannot prove that HNPS in modern English always implies focus on the postposed DP, nor should it. It is merely intended to show that it is necessarily the case that HNPS marks narrow focus in many cases, and pending a more complete corpus study, is likely to be a general strategy for that pragmatic function. It is my hope, since it is notoriously difficult to construct example-discourses to test pragmatic intuitions, and because the intuitions surrounding HNPS are so dependent on the information structure and prosody of the discourse, that I can prove my hypothesis for the pragmatic function of HNPS conclusively at some later point with a broad corpus study (along the lines of what Bies 1996).

[79] Thank you to Josef Fruewald for giving me this example, said by a mutual colleague.

As with the Early New High German data discussed in Bies (1996), a focused constituent does not need to introduce an entirely new entity to the discourse, though it frequently does. This means that the conclusion argued for here, that HNPS is used primarily to indicate (narrow) focus on the DP object by moving it to a specific focus position, is supported by one of the only quantitative studies known on the information status of Heavy NP Shift, Arnold et al (2000). In the terms of Vallduví (1992), the narrow focus is what the hearer is expected to store in a "retrieve-substitute" operation; as Bies (1996: 17) notes, it is "the new information provided by the sentence," i.e. in the context of the rest of the sentence (the "ground" in Vallduví 1992), but it may or may not be completely new information in the context of the preceding discourse. Of course, if a DP does introduce a new entity into the discourse, it is very likely to be narrowly focused in its clause. If the predicate and remainder of the clause also happen to be given in the discourse, then the object represents the information that the speaker does not share with the hearer and wants the hearer to store, and the clause necessarily has narrow focus on the "new" object (Vallduví 1992, following Prince 1985). Arnold et al (2000) found that the postposed object in Heavy NP Shift constructions frequently constitutes information that is new to the discourse, and furthermore, they found that the effect of discourse newness is statistically significant independent of the effect of heaviness. Unfortunately, that study was not able to test finer-grained distinctions of discourse status than "newness" and "givenness". In this way, although they present an important empirical result, Arnold et al (2000) miss the generalization that the syntactic position associated with this type of discourse-newness might be the same as that associated with other

instances of discourse-newness, such as the movement normally (and misleadingly) called "topicalization" (divided into "poset"-Topicalization, i.e. contrastive topicalization, and "Focus Movement" in Prince 1999 and citations therein). However, it is clear from Arnold et al (2000) that prosodic heaviness does not predict the application of Heavy NP Shift alone, and that the "newness" effect which they found is completely expected if Heavy NP Shift is, in fact, a construction of narrow focus on the object.

While canonical topicalization has received much more attention in the literature and it's common to view HNPS as somehow more peripheral, it is interesting to note that HNPS is actually a more common focus strategy in modern English because of its characteristic prosody. Speyer (2005, 2008) showed that the topicalization/Focus Movement of direct objects declined steadily throughout the history of English due to the combination of the loss of V2 and a prosodic constraint which avoids stress clash between a fronted object and an immediately following DP subject. The rate of Focus Movement decreases over time because as V2 declines, the fronted object becomes increasingly adjacent to the following subject, and this resulting string is only prosodically well-formed if the subject is an unstressed pronoun (cf. also Prince 1999: 4, citing Gregory Ward's corpus). By the time of the Early Modern English period (by c. 1500), the rate of direct object topicalization had reached as low as 5.1% of possible cases. In contrast, as I have shown in Table 7 above, the exact same corpus that Speyer used (Kroch, Santorini & Delfs 2004) gave a rate of 14.2% for HNPS at the same time period. To be as precise as possible, I repeated the study above including **only** direct objects, and found the percentage of HNPS to be even slightly higher with direct objects,

15.7% (675/4303 tokens). As Focus Movement/topicalization became increasingly restricted in its possible contexts for application, due to the restriction that it only occur with pronominal subjects, Heavy NP Shift continued to occur at a constant rate. (Of course, HNPS would not suffer from the same prosodic restriction since the subject is far to the left of the focused object in the HNPS construction.) This means that while HNPS cannot necessarily be seen as having actually taken over the ground the Focus Movement previously occupied, it has emerged as the primary movement-based way to indicate focus in modern English.

So to summarize, here are the facts any general analysis of Heavy NP Shift must account for:

1. The construction appears throughout West Germanic (at least), in both Tense-final and Tense-initial varieties
2. In particular, HNPS occurs in Old English Tense-final clauses as well as in modern English, Tense-initial clauses, and does not appear to have changed over time.
3. HNPS is a focus construction in Tense-final and Tense-initial Germanic.
4. The PP/adjunct-scrambling analysis proposed in Kayne (1994: 74) cannot account for all of these at once.
5. HNPS licenses parastic gaps, and so appears to be A'-movement like topicalization.

To this collection of observations I will add one more. The Heavy NP Shift in examples (4-10) above occurs in subordinate clauses. Therefore, for the analysis I am proposing here to work:

6. There must be two specifier positions available high in the phrase structure, one for the focused DP to land in and another for the remnant clause, but they must both be *below* the complementizer position.

7.3 Welsh, Hungarian, and the Split-C System

The analysis of Germanic HNPS in terms of fronting of the focused object, followed by fronting of the remnant clause to a higher specifier position, has a natural statement in the "Split-C" system (Rizzi 1997) argued for on the basis of the Welsh complementizer system in Roberts (2005; and references therein). Roberts argues that the distribution of complementizers, clausal adjuncts, and fronted constituents in Welsh provides solid evidence for a phrase structure in which "CP" is split, and expanded into ForceP, TopicP, FocusP, and FinP. The details of the evidence for the different projections are beyond the scope of this dissertation, but the important point for present purposes is that the lowest complementizer position in Welsh is Fin (as with the complementizer *a*, which follows the clausal adverb and focused DP in example 46 below), the highest is in Force (as with the complementizer *i mai*, which precedes the

clausal adverb and focused DP in example 16 below), and then there are specifier positions for Topic and Focus between the two complementizer positions.

Force **Fin**

(16) Dywedais i mai fel arfer y dynion a fuasai'n gwerthu'r ci
said I MAI as usual the men A would-Asp sell-the dog
"I said that it's as usual the men who would sell the dog."
(Roberts 2005: 32)

The idea that there are two distinct positions for Topic and Focus in the C-domain and that they come in that order (i.e. Spec(TopP) > Spec(FocP)) has independent empirical support from the syntax-information-structure interface in Hungarian. É. Kiss (2002) shows (building on previous work cited therein) that Hungarian has two preverbal positions for fronted constituents: one to which topics move, followed by one to which focused constituents move. É. Kiss defines "topic" as foregrounding "an individual [J.W.: entity]…from among those present in the universe of discourse" (2002: 9) and "focus" (by which É. Kiss means "narrow focus," as opposed to vP/VP focus) as, "a proper subset of the set of contextually or situationally given referents for which the predicate phrase can potentially hold" (2002: 78).[80] Constituents which fill these information-structural roles are fronted to the beginning of the clause in Hungarian, but

[80] These definitions are broadly compatible with the careful treatments of information structure in the literature, notably the "retrieve-substitute" or "tailful focus" of Vallduví (1992: 89), and the discussions of contrastive focus in Prince (1999) and Rooth (1985, 1997), as well as the notion of topic as "Link" in Vallduví (1992). Note also that "topic" may be formalized as the "backwards-looking center" of a clause in Centering Theory (Joshi & Weinstein 1981, Grosz, Joshi, & Weinstein 1995, Walker, Joshi, & Prince 1998, Prince 1999, *inter alia*), though this may end up being too restrictive for either the Hungarian data or the English data I discuss below. I leave a complete comparison of "link" and "backwards-looking center" as an issue for further research.

only in the order Topic > Focus, as shown by the positions of *János* and *Pétert* in (17) below.

(17) János PÉTERT mutatta be Marinak.
John-NOM PETER-ACC introduced VM Mary-to
"As for John, it was PETER that he introduced to Mary."
(É. Kiss 2002: 79)

Hungarian provides clear evidence for two pre-Tense specifier positions, topic and Focus (in that order), which falls neatly in line with Roberts (2005) proposal that the C-system is split into at least two complementizer positions, with Topic and Focus projections below the highest complementizer position in Force. Importantly, Roberts (2005) argues that English and German (and by hypothesis, all of West Germanic) lack the type of complementizer that surfaces low in Welsh (as in example 16 above) in the Fin head position, and rather only contains complementizers that move to, or initially merge in, Force. The opposite system to West Germanic is of course also possible, in which all, or nearly all, complementizers surface in Fin. This is the situation found in modern Irish[81]; Roberts' reanalyzes the arguments made in McCloskey (1996) for the lowering of complementizers in Irish, and shows that the low complementizer position in Irish is really Fin as in the example below.

(18) Is doiche [faoi cheann cúpla lá [go bhféadfaí imeacht]].
Is probable at-the-end-of couple days that could leave
"It's probable that [they] could leave after a couple of days."
McCloskey (1996)

[81] With the possible exception of the Wh-particle *aL* (McCloskey 2001, Roberts 2005: 30).

In this way, Roberts' and Rizzi's (1997) Split-C analysis places the Topic and Focus projections below the surface position of complementizers in English, in exactly the correct relative order suggested by the Hungarian data.

The Split-C analysis, in combination with the independent confirming data on the relative ordering of TopicP and FocusP from Hungarian, predicts the Heavy NP Shift pattern found in Germanic provided it can be understood as a combination of focus- and topic-driven movement. If the postposed DP is focused, as argued above, then we would expect it to move to Spec(FocusP), just as in the canonical English "topicalization" construction (and assuming a Split-C system this is the natural way to handle normal English "topicalization"). The movement of the postposed DP to Spec(FocusP) also explains why HNPS can license parasitic gaps, as in example (2): it is the same type of A'-movement that canonical English topicalization is, and it the DP lands in the same position, Spec(FocusP), so it should have all of the same syntactic properties. In this way, a previously mysterious property of HNPS, that it licenses parasitic gaps, becomes expected and commonplaces under the analysis presented here: HNPS can license parasitic gaps just as any other type of A'-movment can. Once the focused DP is moved to Spec(FocusP), the difference between English DP-"topicalization" and HNPS rests solely on the position of the rest of the clause (the remnant TP). If it is possible that the remnant clause can be marked as a topic by the speaker, which is particularly expected if its predicate is given in prior discourse (cf. e.g. the example in 13 above, the

Pennsylvania German grave markers, and many of Bies' (1996) examples), then it will be fronted to Spec(TopicP). The combination of topic- and focus-marking by moving constituents to dedicated specifiers thereby derives the HNPS pattern from completely independent principles.

7.4 The Fine-Grained Information Structure of HNPS

The specific notion of "topic" that I would like to suggest applies to the fronted remnant TP is the "Link" of Vallduví (1992: 59): "an address pointer in the sense that it directs the hearer to a given address…in the hearer's knowledge-store, under which the information carried by the sentence [J.W.: i.e., the focus] is entered." This notion of topic also corresponds to the presupposition in "focus-presupposition" analyses of focus constructions, as in Prince (1999; and much previous work cited therein); the topic is an "open proposition" which the speaker presupposes to be the true in the world, and it contains a variable which is instantiated by the focus constituent of the sentence (Prince 1999: 6). The open proposition defines a common ground between the speaker and hearer, and then the focus is the information that the speaker signals as being important (or new) to the hearer, and which completes the proposition in terms of the sentential semantics. In Vallduví's system, this topic part of the sentence (his "Ground") is potentially divided into the "Link" and the "Tail", where the Tail specifies some more specific instruction to the hearer about "the exact way in which information is retrieved

and entered under a given address" (Vallduví 1992: 61). The claim here is that HNPS sentences do not have a Tail at all; the entire clause preceding the focused and postposed DP serves as the Link and merely identifies an address for the storage of the DP's information. The remnant TP is fronted without the DP specifically so that it can, as a whole, give instructions to the hearer about how to store the focus *before* the hearer receives the much more important focus information. While more analysis is needed to test this hypothesis carefully, there is one promising indication that the analysis I present here is on the right track: Vallduví (1992: 61) states that one of the few clear attributes of a Tail is that, "it is never marked with prosodic prominence". It is well-known that the constituent preceding the postposed DP in HNPS shows an accent (e.g. on *pockets* in 44 or *wife* in 49), characteristic of the High tone (H*) mark of an intonational phrase boundary (as discussed in Kroch & Pintzuk 1989). This accent is followed by another one on the postposed DP. If we can take the first accent to be marked on the TP preceding the DP, then both of the sentence-level constituents are marked as prosodically prominent. Therefore, HNPS sentences contain no Tail.

The prosodic peak on the constituent immediately preceding the postposed object has another consequence which bears some discussion: it is an indication that HNPS is pragmatically doing something in addition to just marking the postposed DP as narrowly focused and the remnant TP as the topic. So unfortunately, at this point it becomes necessary to slightly complicate the story I've been developing about HNPS thus far. First, as I discussed above, Speyer (2005, 2006) shows that the rate of English topicalization declined in the history of English. In contrast, the rate of HNPS has

remained remarkably stable over time, neither decreasing with canonical topicalization nor increasing to fill a gap left by canonical topicalization. Secondly, while the shifted DP is narrowly focused and the remnant clause is broadly speaking the topic, the presupposition containing the variable to be filled by the focus (in the terminology of e.g. Prince 1999 and previous work), it is not the case that everything within the remnant clause is equally topical (or equally presupposed). This can be seen by applying the question-answer test from (Lenerz 1977: 20; Vallduví 1992: 64, *inter alia*), mentioned above in connection with Bies (1996). Although some of the naturally occurring modern English examples I've discussed involve question-answer pairs, the "question" segment was never the most simple question corresponding to the HNPS answer (though it may have been close); the question contained more, or slightly different, content from merely the predicate of the answer plus a wh-word corresponding to the postposed DP. This is illustrated by the following constructed discourse:[82]

(19) Speaker 1: "Who did you see on Saturday?"

Possible answers from Speaker 2:

a. "I saw my rich uncle from Detroit on Saturday."
b. "On Saturday I saw my rich uncle from Detroit."
c. # "I saw on Saturday my rich uncle from Detroit."

In all three potential answers there is an accent on *my rich uncle from Detroit*, marking it as narrowly focused (i.e. it provides the value for the variable *Who* in the question); the

[82] Thanks to Anthony Kroch for constructing this example and pointing out the relevant contrast.

object has an H* intonational contour and bears the nuclear accent/stress of the clause. This pitch rise is also known as the "primary accent" of the clause in the terms of Ladd (1996), which the accent most usually correlated with a focused constituent (cf. the discussion of primary accent and focus in Ladd 1996: §6.2 and also discussion beginning p.248). In "a" and "b", *my rich uncle from Detroit* is focused in situ, and either placement of *on Saturday* is felicitous. The HNPS sentence in "c", on the other hand, is somewhat awkward in response to the simple question that's paired with the focused object, and it has the feeling (to me, at least) of providing more information than the question is asking in some way.

The prosody of HNPS sentences like the one in "c" is an important clue to the solution of this problem. In the "a" response above, *rich uncle from Detroit* is heavily accented and *on Saturday* is specifically deaccented, being marked with falling intonation: *on Saturday* is clearly a Tail here in Vallduví's system. In "b", *on Saturday* does not have falling intonation, nor is it the most heavily accented constituent in the clause (that is still *my rich uncle*…).[83] In "c", on the other hand, as I mentioned briefly above, both *on Saturday* and *rich uncle from Detroit* are heavily accented to approximately the same degree (the accents are realized specifically on the first syllable of *Saturday* and the last syllable of *Detroit*). This is not unexpected under the syntactic analysis pursued here (which will be elaborated in further detail below), since both the PP

[83] Since there is no falling intonation on *on Saturday* in "b", it is probably not a Tail in that sentence, and it is likely to actually be a topic. I remain agnostic at this point as to whether the PP has simply been adjoined high at the left edge of the clause, or whether it has actually been moved to Spec(TopicP) in that type of sentence. If the latter is true, than I need to mildly revise the statement below that Spec(TopicP) is reserved only for *v*P and TP in English. If PPs can move to Spec(TopicP), then the correct generalization is that Spec(TopicP) is reserved for phrasal categories with the feature [-N] in the sense of Chomsky (1970).

and the DP appear at a major clause boundary, TP and ForceP (i.e. the analog of CP in the split-C phrase structure), respectively. Since they each occur in their own clause, in a sense, both constituents are able to bear the primary, or nuclear, accent of a clause. Of course, the fact that the PP **can** bear a nuclear accent (in addition to the object) does not mean that it must do so, so the fact that it does is an indication that the PP also plays an important role in the information structure of the class. After all, it is perfectly well-formed for a clause-final PP to be completely destressed with falling intonation, as in the "a" response above.

The double accent in HNPS clauses, on the DP and on the final constituent of the remnant TP in Spec(TopicP), shows that HNPS is in fact a double-focus construction, even though only one constituent is actually moved to Spec(FocusP) in overt syntax. The study presented here has already established that the postposed object is narrowly focused, but I would suggest that both accents are marking narrow focus in the HNPS construction. This is why the questions paired with HNPS answers in naturally occurring speech (as in the modern English examples cited above) are rarely of the simplest form, resembling the HNPS answer in all but their questioning of the narrowly focused DP, and why a HNPS answer to this type of simple paired question sounds awkward. The HNPS sentence is in fact providing more information by focusing the adjunct immediately preceding the object as well as the object itself. The HNPS above is much closer to answering a question like, "When did you see WHO?" The adjunct provides additional information by restricting the meaning of the *v*P inside the remnant TP; the set of models (i.e. possible worlds) under which the predicate in the *v*P is true is now restricted to the

subset under which the function represented by the adjunct is also true. In this way, while the basic predicate (aside from the adjunct) and the rest of the remnant TP is a topic in the discourse, the adjunct is a type of contrastive focus, picking out and emphasizing a particular subset of the models in which the predicate holds of the object.

The two naturally occurring modern English examples below demonstrate the double-focus nature of HNPS especially well.

(20) We have found we can enjoy, sober, every good thing we enjoyed while drinking.
(Anonymous 2009: 42; see reference section for more information)

(21) This refrigerator is cleaned out every Friday morning. All unsealed, unlabeled food will be discarded.

Please label with your name anything that you wish to keep!
(Sign posted on the common refrigerator at IRCS, at Penn[84])

Both of these examples have two things in common: first, that both the adjunct and the postposed DP must be accented, and secondly, that the basic predicate of the HNPS clause (minus from the adjunct) is topical in context (i.e. presupposed by the speaker), but the adjunct itself is not. The commas in the first example are particularly helpful,. because they show that the author of the sentence clearly intended both the adjunct and the "shifted" object to be set off from other elements prosodically and accented. This sentence also occurs in a chapter which is all about enjoying various things, so the predicate *enjoy* is shown by context to be topical. The adjunct, *sober*, is not topical, however; the point of the sentence is that things can be enjoyed either sober or

[84] Thanks again to Josef Fruewald for this example.

intoxicated. A paraphrase of the sentence is: there is something, x, which can be enjoyed in one state of mind, y, and importantly, x = "every good thing…" (in contrast to other possible DPs, e.g. "only **some** good things…") and y = "sober" (in contrast to a set of other possible states of mind, e.g. intoxicated). The second HNPS example above actually contains the predicate *label* in the preceding sentence, so that verb is clearly topical. The next sentence goes on to provide the reader with two specific pieces of information about *labeling*: the label must contain your name (in contrast to other marks which might not be as useful, in case you graduate and leave old food or some such), and you only have to label food that you want to keep (in contrast to other food, which you are free to leave to be thrown out on Friday).

Following the account of focus in Rooth (1985, 1992, 1997), both focused constituents in a HNPS construction bear a syntactic [+ Focus] feature. This is rather a nice proposal, since according to Minimalist assumptions, this is a necessity of the theory of movement: in order for anything to move to Spec(FocusP), as I am arguing, there must be some [Focus] feature for the Focus head to probe for and trigger movement.[85] (Similarly, the TP in a HNPS sentence must be marked with a [Topic] feature.) However, in the restrictive version of the Split-C hypothesis adopted here, there is only one FocusP projection per clause. This being the case, as the Focus head probes down the tree for a [Focus] feature, either the focused adjunct or the focused DP may be

[85] In the theory I have developed in this dissertation, adjunction in an exception to this in being truly optional. Scrambling does not require features on the scrambled elements, and for this reason, it must be globally constrained by the Conservation of C-Command, as I have already discussed. However, following the Split-C framework, focus- and topic- related movements land in dedicated specifier positions, and therefore they are not adjunctions and must be feature-driven. The fact that focus and topic movement are feature-driven also has the consequence that there must be covert (LF) topic and focus movements as well in those cases where these specifiers are not filled on the surface.

attracted to Spec(FocusP), but not both simultaneously.[86] Assuming the DP is attracted, it fills Spec(FocusP), and then the TP marked with [Topic] is attracted to Spec(TopicP) and the HNPS construction results. This allows the topic and focused object to be marked syntactically at the same time as aligning the focused adjunct with the nuclear (primary) accent at the right edge of its TP so that it can most easily be marked for focus prosodically. (It may in fact be the case that the syntactic marking of the two foci and the topic is preferred, and purely prosodic marking of these pragmatic roles is a last resort, but further research is required to determine if this is actually true.)

Importantly, this derivation shows that the syntactic apparatus is **not** isomorphic with the information structural component of the linguistic system, and this is exactly the correct prediction to make. As Prince (1999 and previous work) shows, English and Yiddish topicalization demonstrates the autonomy of syntax from information structure in two ways: the same construction serves more than one pragmatic function, and there is never a situation in which it is obligatory; there is always the possibility of conveying the same information structure in some other way (e.g. with prosody alone). HNPS is another such example of non-isomorphism, in that the syntax goes as far as it can to mark each of the foci and the topic, but independent constraints on the syntax prevent it from marking the second focus syntactically (i.e. there's no place to put it![87]). The hypothesis of the autonomy of syntax from other grammatical components predicts that such

[86] Or, if the basic order is Object > Adjunct, then according to Kayne (1994), the object c-commands the adjunct in some way, and it is only the object that will be attracted to Spec(FocusP) in this case since it will be encountered by the probe earlier in the tree than the adjunct will. In either case, the HNPS construction can be the result.

[87] At least, in overt syntax English doesn't appear to allow double specifiers of FocusP, which would result in duuble-topicalizations without HNPS.

mismatches should occur. If we never find this kind of problem, then we should really begin to question the entire idea of a there is such a thing as a purely syntactic constraint or an arbitrary syntactic requirement. Happily, we do find independent syntactic constraints asserting themselves, and I would suggest that the interface with information structure is a good place to seek them out.

Hungarian and West Germanic are maximally similar under this view: they both have a Split-C system, as in Celtic, with a TopicP and a FocusP, and English and Hungarian clauses can potentially move both a Topic to Spec(TopicP) and an argument to Spec(FocusP) simultaneously. I have already illustrated this for modern English in the tree shown in (3) above, which could easily have served as the derivation of the Hungarian sentence in (17) if both the topic and focus were DPs. This analysis is made much more plausible by the Split-C analysis that Roberts (2005) is led to for Welsh, because it not only places the Topic and Focus projections in the appropriate relative positions for West Germanic, but it also places Germanic complementizers above those two projections, in the highest complementizer position that the Celtic system shows. This leads to the correct prediction that Heavy NP Shift can occur in subordinate clauses in West Germanic, as it does in a number of examples above and in the modern English[88] examples below (I have given the whole sentences below, but the relevant subordinate clause is bracketed).

---

[88] These examples are from 18th and 19th century English only because the most accurate modern English, fully parsed corpus I currently have access to is Kroch & Santorini (*forthcoming*), which only contains texts up to 1910. Such examples are also grammatical in 21st century English (possibly modulo a few minor archaisms), as any native English speaker reading this thesis should be able to confirm.

(22) I also recollected [that I had mentioned to my wife my intending to be home this week].
(James Boswell's diary 1776-1778, *Boswell in extremes*; BOSWELL-1776,45.331 in Kroch & Santorini *forthcoming*)

(23) What does Sir B. Frere propose to do with regard to ships suspected of carrying slaves but carrying the flags of countries [which have not conceded to us the right of search]?
(*The political correspondence of Mr. Gladstone and Lord Granville, 1868-1876*; GLADSTONE-1873,2,382.212 in Kroch & Santorini *forthcoming*)

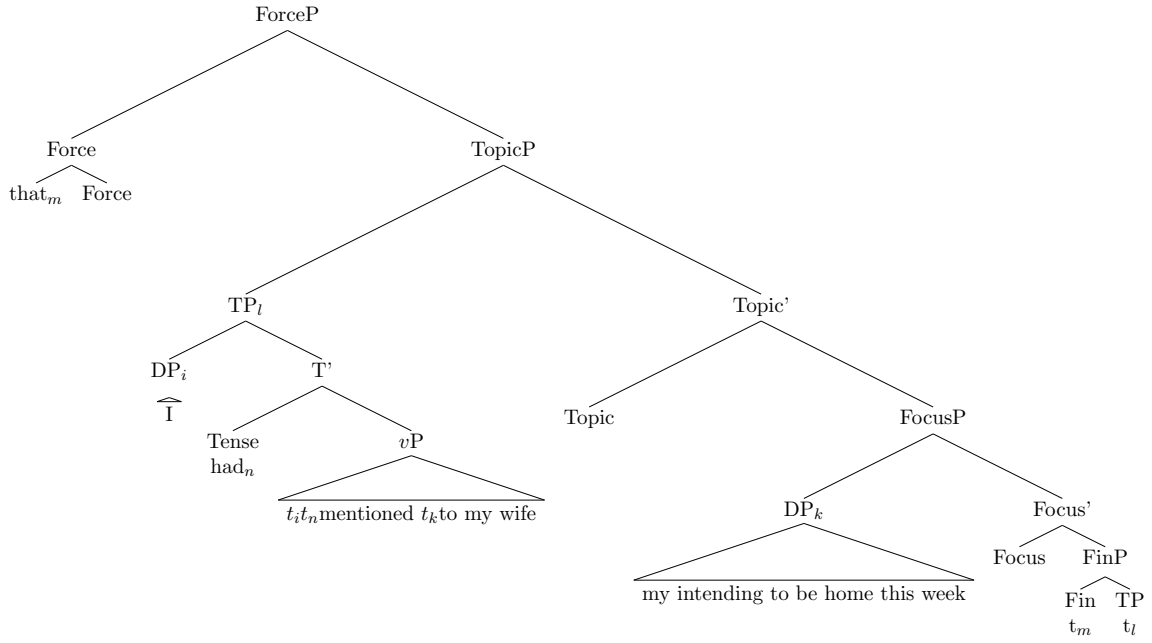
(24) It is a marvel to us, how at least fidelity on the wife's side could become to such an extent a heathen virtue.
(Edward B. Pusey's *Lenten sermons, preached chiefly to young men at the universities, between A.D. 1858-1874*; PUSEY-186X,301.297 in Kroch & Santorini *forthcoming*)

(25) The letter of the former is somewhat imprudent, [upon which I will communicate to him a piece of my mind];
(*Selections from the dispatches and general orders of Field Marshal the Duke of Wellington*; date: 1815; WELLESLEY-1815,833.7 in Kroch & Santorini *forthcoming*)

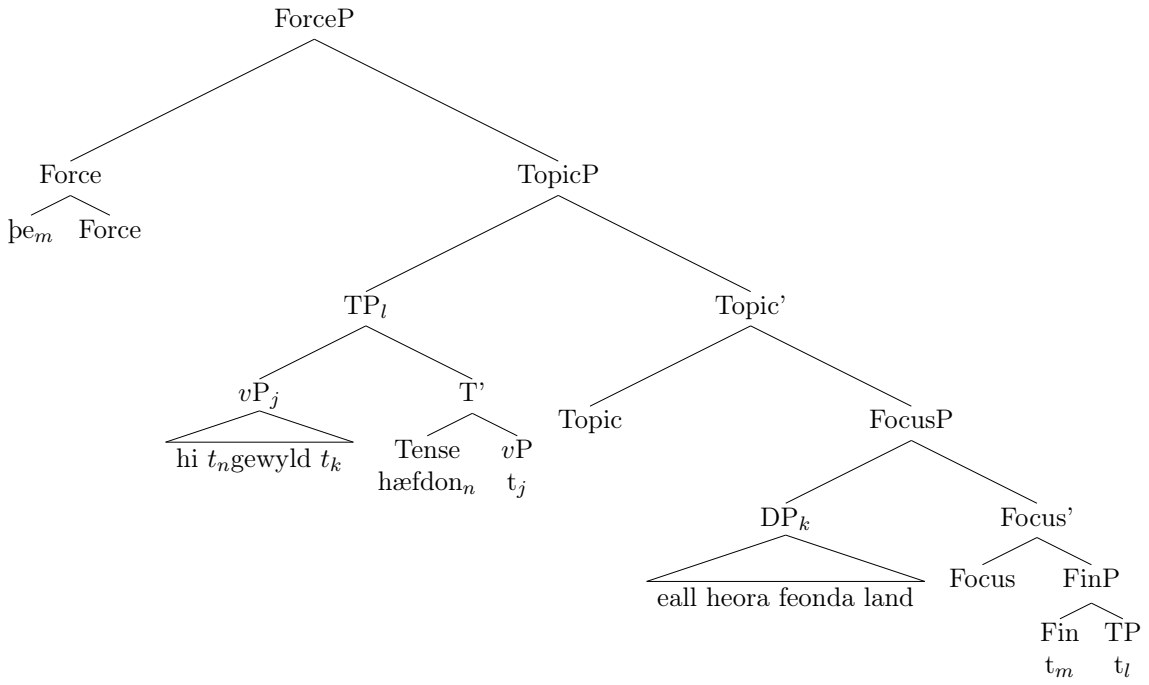
The relative clause examples in (51)-(53) are particularly important because they make clear that TopicP and FocusP are below the complementizer position in Germanic. For this reason, the fact that both the object and the remnant clause have been fronted does not interfere with the wh-movement of *which* out of the TP within the relative clause. Under standard assumptions, the landing site of the wh-phrase is just above the complementizer, in a position which I take to be Spec(ForceP) in the Split-C phrase structure. And of course, there is no reason why this position should be restricted to topic or focus constituents.

At this point, it is possible to give the full structure for a unitary analysis of HNPS across Germanic (in 26 below), which I illustrate with the Old English subordinate clause from (5) and the modern English subordinate clause in (22).[89]

(26)



[89] In these trees I show the Fin-to-Force movement of the complementizer, as Roberts (2005) argues for Germanic complementizers in the complement of non-bridge verbs. I did not show cyclic head movement of Fin through Focus and Topic as well merely for ease of presentation, since this detail has no bearing on the current discussion. I do not take any stand on whether the complementizer also incorporates with Focus and Topic on its way up the tree.



In both structures, the complementizer has head-moved moved from Fin to Force, the TP has been fronted as a topic to Spec(TopicP), and the object has been focused and fronted to Spec(FocusP). There is no difference between the structure below and the structure of modern English HNPS above the level of the TP. In this way, the analysis here correctly predicts that HNPS did not change at all during the course of the history of English, though the internal structure of TP obviously did.

It's necessary at this stage of the discussion to make a few brief remarks on the fact that focus movement occurs within subordinate clauses according to the analysis I've pursued here. As discussed in Rooth (1992) and Drubig (1994), who address the issue in different ways, although constituents within subordinate clauses, focus must be interpreted at the level of the matrix clause. However one addresses the problem of how

to interpret embedded foci as if they were not embedded, (and a full discussion is beyond the scope of this dissertation), one might ask why there is overt focus movement to the specifier of a FocusP projection within subordinate clauses. The answer is, as far as I can say at present, there is no answer other than that syntactic operations are autonomous from other components of grammar and frequently independent from each other as well. It is simply a fact about syntax (at least Germanic and Celtic) that every clause has a FocusP projection by virtue of its being a clause (following Rizzi 1997, Roberts 2005, *inter alia*), just as it had a CP node under the classical non-split-C phrase structure. In the most simple case the DP is attracted to the embedded Spec(FocusP), where it is then further attracted to te matrix Spec(FocusP) either in overt syntax, resulting in cyclic topicalization, or covertly by LF movement, resulting in embedded HNPS on the surface. However, this cannot occur if the subordinate clause happens to be an island, as it is in the relative clause examples above. This is the problem for which Drubig (1994) proposes the pied-piping of the whole subordinate clause at LF to a position in the matrix clause. Some solution along these lines could be adopted here, with the whole ForceP being pied-piped to the matrix Spec(FocusP) at LF. Indeed, a solution like this is already needed for the HNPS construction since the focused adjunct cannot move to the embedded Spec(FocusP) at all; there must be some mechanism for interpreting more than one focus in an extraction island. But however the problem of islands is addressed, it does not preclude movement to the embedded Spec(FocusP) within the island, since there is still a Focus head there (by the Split-C hypothesis) and that will attract the closest [+ Focus] constituent in a purely dumb and mechanical way,

without reference to the fact that the island prevents further cyclic movement into the matrix clause. In this way, the Minimalist theory of feature-driven movement predicts that focus movement, and hence HNPS, should be possible within an island, even if it does not always automatically contribute to the matrix evaluation of focus.

7.5 TopicP and *v*P-Topicalization in Modern English

As I said above, Hungarian and English are maximally similar under the account presented here. But of course, there is a difference between Hungarian and West Germanic: the only difference between the two Split-C systems is the nature of TopicP. In Hungarian, Spec(TopicP) can host a topic of nearly any syntactic category (for details I refer the reader to É. Kiss 2002 and references therein). But apparently languages can differ on the types of objects which can be fronted to Spec(TopicP), just as languages can specify features of the type of element that can be moved to Spec(TP) to serve as a subject. And whatever parameter determines which type of phrase can be fronted to fill Spec(TopicP), Hungarian falls on the most permissive side of that parameter and English falls on the other side: English, and West Germanic generally, can only truly "topicalize", in this more precise sense, the phrases TP and vP (maximal categories along the extended projection of the verb). This approach makes a very clear prediction beyond HNPS: first, that TPs can be fronted to Spec(TopicP) even when nothing is focused. Unfortunately, this is a difficult prediction to test, since fronting a TP without moving anything else out

will be string-vacuous movement. However, it is possible that this is in fact the structure of some sentences that do not have a Tail (a type discussed in Vallduví 1992) and which do not front their focused element for any number of reasons (e.g., the focus is on a head rather than a phrase, so it cannot be moved to a specifier position). A more interesting prediction is that vP could be fronted to Spec(TopicP), with or without accompanying Focus Movement, and it should have effects which are distinct from the movement of categories to Spec(FocusP).

It does. vP-topicalization does not pattern with other types of English topicalization, either prosodically or information-structurally. Prince (1999) distinguishes two different subtypes of the modern English fronting construction usually termed "topicalization", and vP-topicalization does not fall neatly into either category. Prosodically, it involves an accent on the vP (on its head, the lexical verb), as well as an accent within the following clause on the auxiliary (or periphrastic "do"), as in the example below:

(27) Mary wants to finish grad school, and finish grad school she did!

This makes vP-topicalization prosodically very unlike Prince (1999)'s "Focus Movement", which does not include a second accent within the remnant clause, and most like Prince (1999)'s "poset"-Topicalization (analogous to "contrastive topicalization/focus" in many other discussions of focus). However, unlike the "poset" type of contrastive topicalization, the fronted vP does not normally name an entity that

contrasts with some other set of entities present in the discourse.[90] In fact, it differs from all focus constructions in arguably providing no new information to the listener in its fronted constituent. As opposed to the narrow focus (of which both types of topicalization discussed in Prince 1999 are subtypes), which usually presents a new entity into the discourse, the predicate that the fronted vP refers to *must* be given in the discourse or salient in the mind of the hearer in some other way; indeed, vP-topicalization is most felicitous in modern English when it repeats a predicate that was just mentioned in the immediately preceding clause. Furthermore, if anything in the clause containing the vP-topicalization is narrowly or contrastively focused, it is the Tense element or Neg,[91] not the vP. This is shown in (28) below, which is followed by an example showing that vP-topicalization is infelicitous without some kind of contrast in the Tense element between the vP-topicalization clause and the preceding one.

(28) a. I said I would quit my job if Bush was elected, and quit my job I díd!
b. Either John can do his work, or do his work he cannót.

(29) # John might finish his paper tomorrow, and finish his paper he might!

---

[90] To be clear, I am certainly **not** claiming that it's impossible to front a vP in modern English and contrastively focus it. This would be an instance of one of the two fronting constructions that Prince (1999) discussed. I am only claiming that there is another fronting construction, with a different pragmatic and prosodic profile, which is not available to other clausal constituents, e.g. DP objects.

[91] The examples where the focus in the second clause is only on Neg sound somewhat stilted in modern English, though they did occur in Early Modern English:

ii. and grawnte yt he wold not yn no wyse.
and grant it he would not in no wise
"And he wouldn't grant it [a loan] at all."
(*The autobiography of Thomas Mowntayne*, in Kroch, Santorini & Delfs 2004; date: 1555)

Perhaps this is really only a literary device, and the focus really needs to be on Tense in natural speech, or perhaps a subtle change took place since EME times with regard to this construction.

Even if we accepted the prosody of vP-topicalization as anomalous, it would still not fit in terms of information structure with the other category of topicalization analyzed in Prince (1999). The other type, "Focus Movement", fronts a constituent that identifies the particular value of some attribute that another entity in the clause already has (and is known to have by the hearer). The only possible candidate for the attribute-carrying entity in this construction is the subject, and judging from examples of the type that those in (28)-(29) illustrate, the fronted vP does not name any new attribute of the subject's. Again, if there is any new attribute of the subject that is named by the clause in which the vP-topicalization takes place, it is identified by the Tense element in the clause.

However, the preceding observations are entirely expected if the fronted vP is a Topic (or a "Link", in Valduví 1992). The only difference between HNPS and vP-topicalization is that it is the vP that is topical and fronted to Spec(TopicP), rather than the TP. But if that's true, then why does vP-topicalization appear to be so much more restricted in its felicitous contexts than HNPS? The answer is that having the vP be topical is not the only condition for vP-topicalization: as I mentioned before, vP-topicalization makes the vP the topic *and* makes the Tense element the focus of the clause. In fact, vP-topicalization seems to be specialized to this pragmatic function of focusing Tense, as demonstrated by the contrast in (28) and (29) above. There is a two-fold reason for this: first, there is no way to focus the Tense element by fronting it since it is a head rather than an XP, so the only option is to mark the focus prosodically. Topicalizing the vP places Tense at the end of the sentence and the end of the sentence's

intonational phrase. This prosodically lines the Tense element up with the default phrase-final High boundary tone (the "tonic stress" of the clause), and so now the normally unstressed auxiliary in Tense (and most purely functional elements are unstressed by default) is easy to stress and mark as the focus. HNPS makes the TP topical and focuses a DP object, while vP-topicalization fronts the vP and focuses Tense. Presumably, contexts which have the combination of factors that the vP is given in the discourse and Tense is the natural focus are rare enough that vP-topicalization is highly restricted. Indeed, a quick corpus study bears this out: a search of the PPCEME (Kroch, Santorini & Delfs 2004) shows that only 12 out of 22656, or 0.05%, of finite matrix clauses with auxiliaries contain vP-topicalization (compared with 14% for HNPS, as I mentioned above).

Since vP-topicalization does not necessarily move anything to Spec(FocusP), this analysis predicts that it should be possible to front something to Spec(FocusP) at the same time as the vP is fronted to Spec(TopicP), just as in canonical HNPS. Indeed this is true, as in examples such as the ones below from Culicover & Rochement (1990).

(30) Everyone said that John would give to Mary all of the money that he won at the track, and give to Mary all of the money that he won at the track he did!

(31) John was told to buy for Mary every book he could find, and buy for Mary every book he could find he did!
(Culicover & Rochemont 1990: 119)

Under the analysis argued for here, the sentence after the conjunction in (30) contains true topicalization of the vP, movement to Spec(TopicP), as well as movement of the DP

all of the money… to Spec(ForceP), placing it below the fronted vP but above the remainder of the clause; the relative order of the two fronted elements is predicted: Topic, then Focus. Thus, the apparent constituency of the DP and (the rest of the) vP which Culicover & Rochemont (1990) are arguing for is actually illusory. Incidentally, Culicover & Rochemont do not happen to remark on the fact that these examples sound somewhat awkward; but this is not really surprising. Since the pragmatic context for vP-topicalization is already quite restricted, the context for this type of sentence is even more restricted: such sentences can only be a double-focus construction, focusing the fronted object and Tense at the same time, while also asserting that the vP is topical.

Finally, the analysis of HNPS presented here is importantly different from the analysis of verb (projection) raising (VPR) in Tense-final West Germanic that I presented above in section 6.5, contra the tradition of Zaenen (1979), Kroch & Santorini (1991), among others, who analyze VPR as a type of rightward extraposition transformation. In part, the analysis of HNPS in this section was inspired by the fact that the approach I argued for above in 6.5 for VPR is not extendable to HNPS; when the lowest vP in the structure is pied-piped to the left of the nonfinite verb, as I showed in 6.5, there is no possibility for a DP to be stranded to the right of the nonfinite verb because there is no structure left to the right of the nonfinite verb for the DP to attach to. Of course, there is structure available for the DP to attach to as the vP cyclically moves through the rest of the tree in a verb-raising language like Dutch or West Flemish.

While the incompatibility of these two analyses may appear to be a drawback at first glance, the non-unitary analysis of VPR and HNPS makes an important prediction.

As I have mentioned above a number of times, the prosody of HNPS involves an intonational break preceding the postposed DP, and Kroch & Pintzuk (1989) showed that this was true of Old English as well (and by hypothesis, all of Tense-final West Germanic which has or had the possibility of HNPS). If VPR were a similar type of extraposition process, with the nonfinite vP(s) moving rightward to attach to the right of the clause (as in the classical analysis of Heavy NP Shift), then the clause boundary would be marked by the position of the finite verb in a VPR clause like the one below.

(32) as hulle daar moet goeie onderwys gee
if they there must good education give
'if they must provide a good education there'
(Afrikaans example from Robbers 1997: 76)

In other words, the position of the finite verb *moet* in example (32) would be identical to the position immediately to the left of the HNPS-object, whether that is an adjunct, as in the modern English examples, or indeed the finite verb in Old English examples such as (36) above or those discussed in Kroch & Pintzuk (1989). If verb (projection) raising were extraposition of the non-finite verb (and its potentially its complement), then we would naturally assume it would have the same prosodic structure as heavy NP shift, all things being equal, particularly if the raised projection included just an object.

However, the prosody is not parallel at all. In fact, there is no intonational phrase boundary or accent on the finite verb in VPR structures like the one in example (60) above (Theresa Biberauer, p.c.). If we assume a reasonably straightforward correspondence of major clause boundaries to major intonation boundaries, then the

difference in prosody between VPR and HNPS shows us that the clause boundary is **not** immediately to the left of the DP *goeie onderwys* in (32), at the position the finite verb. Rather, the clause boundary is at the intonational boundary, to the right of the DP at the position of the nonfinite verb *gee*. Thus, the analysis of HNPS I presented here and the analysis of VPR I presented above predict the correct prosody for the two constructions without any ad hoc stipulations about the prosody-syntax interface, unlike any approach that treats HNPS and VPR as related constructions.

7.6 Summary and Conclusions

In this section I have shown that not only is an antisymmetric account of Heavy NP Shift possible, but also that it can provide a unitary analysis which covers the data across historical and contemporary West Germanic. Importantly, it does so without inventing **any new theoretical machinery**: the idea of a more articulated CP structure, "Split-C", has solid, independent empirical support from Celtic (Roberts 2005), and the TopicP projections with a FocusP complement is independently confirmed by the Hungarian data (É. Kiss 2002). Furthermore, the analysis makes correct predictions about the interaction between HNPS and a number of other constructions, while also providing a straightforward account of the mapping between HNPS's syntax and its prosody and information structure.

Chapter 8

Reconstruction and the Reverse Webelhuth's Paradox

8.1 Introduction: "A-" vs. "A'-" Scrambling

In this chapter I briefly address the question of whether there are two types of scrambling, A- and A'-, or only one. Mahajan (1990) and others following his study (including Lee & Santorini 1994) have argued on the basis of binding data that there is A'-scrambling, which moves constituents leftward across the subject, and A-scrambling, which moves constituents more locally. This conclusion is reached based on the observation that objects scrambled leftward across the subject generally reconstruct to a lower position for the purposes of binding, and objects scrambled to positions below the subject do not reconstruct. In the context of this dissertation, the question becomes: why should the same constraint, the GHC/Conservation of C-Command, apply to two apparently different types of movement?

The answer I will argue for here is that scrambling is not two types of movement, but only one. It is a type of adjunction, following Saito (1985), Webelhuth (1989), and

Lee (1993), but it is an unusual type being adjunction of an element with an internal source: while most adjunction by external merge, scrambling is internal merge, i.e. movement. As such, it is a type of movement which is distinct from both A- and A'-movement (contra Mahajan 1990, Webelhuth 1989, Lee 1993, and Lee & Santorini 1994), which are both movement to dedicated specifier positions. (In this I follow the suggestion in Chomsky 1993 that there may be three types of movement: A, A', and adjunction). Internal adjunction, on the other hand, is a general movement with many possible targets in the clausal structure, which extends maximal projections beyond the specifier level.

In particular, I argue that the data on reconstruction for binding are not a barrier to a unitary analysis of scrambling, as they do not show that non-local scrambling is A'-movement when considered carefully, nor do they support the conclusion that scrambling is A-movement. They do show, however, that scrambling and Scandinavian object shift behave alike with respect to binding. And finally, I show that an effect which I've termed the "Reverse Webelhuth's Paradox" demonstrates conclusively that scrambling/object shift past the subject is not A'-movement.

8.2 Reconstruction does not necessitate 2 types of scrambling/object shift

As I pointed out briefly above in Chapter 3, when the subject of a clause is a potential binder for some other argument, scrambling the argument to the left of the

subject does not change the binding relations; scrambling shows reconstruction in this configuration. Frank, Lee, & Rambow (1996), Lee & Santorini (1994), and Lee (1993) established this fact for scrambling in German and Korean on the basis of both quantifier binding and anaphor binding. A very simple example in German is the obligatory reconstruction of the reflexive, *sich*, in the example below, which would constitute a Principle A violation if it were interpreted in its surface position.

(1) Gestern hat sich Johann rasiert.
Yesterday has SELF John shaved.
"Yesterday, John shaved himself."

Similarly, the second Korean example below in (2b) shows obligatory reconstruction of the scrambled object, *Minswu-uy emma-lul* ("Minswu's mother"). The first example in (2) shows a clear Principle C violation, with the subject *ku-ka* illicitly binding *Minswu* inside the object. The "b" example shows that scrambling the object above the subject does not repair the violation, showing that the object must reconstruct in the position of the trace below.

(2) a. * ku_i-ka $Minswu_i$-uy emma-lul coahanta
he_i-NOM $Minswu_i$-GEN mother-ACC like
"He likes Minswu's mother"

b. * Minswu-uy emma-lul ku-ka t coahanta
$Minswu_i$-GEN mother-ACC he_i-NOM like
(Lee 1993: 28)

To these facts I have added Yiddish scrambling under V-to-C movement and Swedish

"long object shift" (which I take to be a sub-case of scrambling), both of which show reconstruction of reflexive objects that have been scrambled past the subject. In fact, Holmberg (1986) and Josefsson (1992) both report that reflexive objects are one of the most likely types of objects to be moved to the left of the subject (Josefsson also shows that this phenomenon is in fact more widespread with other types of pronominal objects than Holmberg originally claimed). This is shown in the Yiddish and Swedish examples below, repeated from Chapter 3.

(3) farvos zoln **zikh** yidn glat krign?
why should **SELF** Jews in-general fight
"Why should Jews always fight amongst themselves?!"
(Perets Hirshbeyn, *Grine Felder*, token 103, date: 1910)

(4) Igår lade **sig** mamma tidigt.
yesterday laid **SELF** mother early
"Yesterday mother went to bed early."
(Josefsson 1992: 67)

(5) Klarar **sig** barnen på egen hand?
manage **SELF** children-the on own hand
"Do the children manage on their own?"
(Hellan & Platzack 1995: 58)

However, contrary to what is often assumed (cf. e.g. Webelhuth 1989, Mahajan 1990, Lee & Santorini 1994, and others following those original observations), scrambling past the subject does not always trigger reconstruction. Lee (1993) and Frank, Lee & Rambow (1996) show conclusively that scrambling of an object past the subject does not show reconstruction effects unless the subject is a potential binder for the object. This conclusion was reached based on data from Korean such as the examples

below.

(6) a. Minswu-uy tongsayng$_i$ ku$_i$-eykey sacin-ul poyecwuessta.
Minswu-GEN brother$_i$-NOM him$_i$-DAT picture-ACC showed
"Minswu's brother$_i$ showed him$_i$ a picture."

b. * ku$_i$-eykey Minswu-uy tongsayng$_i$ t$_i$ sacin-ul poyecwuessta.
him$_i$-DAT Minswu-GEN brother$_i$-NOM picture-ACC showed
(Lee 1993: 28)

The sentence in (6a) is the unscrambled order and (6b) shows the object, *ku-eykey* ("him") scrambled leftward past the subject, *Minswu-uy tongsayng* ("Minswu's brother"). The latter order is ungrammatical creates a Principle C violation due to the fact that "him" now binds the genitive "Minswu" inside the subject. Clearly no reconstruction has taken place or can take place in this context, specifically because the subject is not coindexed with the object and thus is not a potential binder for the object. It is plain from this contrast that scrambling past the subject does not categorically display the A'-property of reconstruction, and therefore, one cannot argue from the reconstruction facts that this type of scrambling is true A'-movement.

Japanese shows the same contrasts as Korean with respect to Principle C and scrambling past the subject, as shown in the examples below. In these examples I have chosen "John's sister" as the subject, making it even clearer that the subject is not a potential binder for the masculine object "him".

(7) a. John$_i$-no imooto-ga kare$_i$-ni gohan-o ageta
John$_i$-GEN sister-NOM him$_i$-DAT food-ACC gave
"John's sister gave him food."

b. * kare$_i$-ni John$_i$-no imooto-ga gohan-o ageta
him$_i$-DAT John$_i$-GEN sister-NOM food-ACC gave

Just to underline this point, the sentences below in (8) show that the subject is not a potential binder for the object in (7); a reflexive pronoun object is not acceptable in either scrambled or non-scrambled position when it is coindexed with "John".

(8) a. * zibunzisin$_i$-o John$_i$-no imooto-ga hihansita
SELF$_i$-ACC John$_i$-GEN sister-NOM criticized
Intended: "John's sister criticized himself"

b. * John$_i$-no imooto-ga zibunzisin$_i$-o hihansita
John$_i$-GEN sister-NOM SELF$_i$-ACC criticized

Again, the "b" example in (7) above shows that reconstruction of the scrambled object pronoun is not a possibility, even though it is scrambled past the subject, because the subject is not a potential binder for the object. And of course, both the scrambled and non-scrambled configurations are possible in Japanese if the genitive DP within the subject is not an R-expression. Below I have replaced *John-no* with the genitive pronoun, *kare-no*, which may be bound or unbound. No Principle C violation arisies, and both orders become possible as expected.

(9) a. kare$_i$-no imooto-ga kare$_i$-ni gohan-o ageta
his$_i$-GEN sister-NOM him$_i$-DAT food-ACC gave
"His sister gave him food."

b. kare$_i$-ni kare$_i$-no imooto-ga gohan-o ageta
him$_i$ –DAT his$_i$-GEN sister-NOM food-ACC gave

Japanese also demonstrates the lack of reconstruction for scrambling past a non-binder subject with a striking Principle A effect. A scrambled object in Japanese can bind a reflexive or reciprocal in the specifier of a subject, provided that the subject is not itself a possible binder for the object, as shown in (10).

(10) a. * zibunzisin$_i$-no imooto-ga John$_i$-ni gohan-o ageta
SELF$_i$-GEN sister-NOM John$_i$-DAT food-ACC gave
Intended: "Himself's sister gave John food."

b. ? John$_i$-ni zibunzisin$_i$-no imooto-ga gohan-o ageta
John$_i$-DAT SELF$_i$-GEN sister-NOM food-ACC gave
"Himself's sister gave John food."[92]
(Satoshi Nambu p.c., for data 7-10)

This effect was also noted by Saito (1992) for reciprocal binding, based on data like the following:

(11) Karera$_i$-o otagai$_i$-no sensei-ga t$_i$ hihansita
Them$_i$-ACC each other$_i$-GEN teacher-NOM criticized
"Each other's teachers criticized them"
(Saito 1992)

In (10b) and (11), the object in its surface, scrambled position is able to bind the reflexive or reciprocal within the subject. This would not be possible if reconstruction took place; as (10a) shows, if the object remains in its underlying position the reflexive in the subject's specifier is unbound and a Principle A violation produces ungrammaticality.

[92] The question mark on the "b" sentence refers to the following: while the Japanese speakers I consulted found the contrast between the scrambled and unscrambled orders here to be robust, they did not agree on whether the "b" example was entirely grammatical. Satoshi Tomioka (p.c.) suggested that the subtlety of the judgment is due to the fact that the compound reflexive "zibunzisin" has a strong tendency to be subject-oriented, though it is not completely unable to be bound by non-subjects.

Again, an object scrambled past the subject cannot reconstruct unless the subject is a potential binder for the object.

As expected given the previous results in Lee & Santorini (1994) and Frank, Lee, & Rambow (1996), German scrambling shows the same effects as the Japanese and Korean data above. The same lack of reconstruction when the subject is not a potential binder can be seen in the German "b" sentences below. *John* is coindexed with the pronoun *ihm* ("him") below, but embedded within a subject ("John's sister") which is not a potential binder for *ihm*. When *ihm* is scrambled past the subject, it binds into the subject and results in a Condition C violation by binding *John*. No reconstruction of the scrambled pronoun takes place, even though it has undergone scrambling past the subject. (The same contrast was reported for parallel German subordinate clauses in Frank, Lee, & Rambow 1996: 5.)

(12) a. Gestern hat die Schwester von John$_i$ ihm$_i$ ein Bild gezeigt.
Yesterday has the sister of John$_i$ him$_i$ a picture showed.
"Yesterday, John's sister showed him a picture."

b. * Gestern hat ihm$_i$ die Schwester von John$_i$ ein Bild gezeigt.
Yesterday has him$_i$ the sister of John$_i$ a picture showed.

(13) a. Gestern hat Johns$_i$ Schwester ihm$_i$ ein Bild gezeigt.
Yesterday has John$_i$'s sister him$_i$ a picture showed.
"Yesterday, John's sister showed him a picture."

b. * Gestern hat ihm$_i$ Johns$_i$ Schwester ein Bild gezeigt.
Yesterday has him$_i$ John$_i$'s sister a picture showed.
(Lucas Champollion, Beatrice Santorini[93], p.c.)

[93] For reasons that are unclear, Beatrice Santorini found the starred examples less egregious than Lucas Champollion did.

Much of the literature assumes that Scandinavian object shift and German scrambling are different phenomena, contrary to the position I have taken throughout this dissertation. This may account for the fact that it has not yet been noticed, to my knowledge, that object shift behaves identically to German, Japanese, and Korean scrambling with respect to these binding facts, once appropriate examples are constructed. As I mentioned above, Swedish object shift of a reflexive pronoun past a potential subject binder results in reconstruction, binding the reflexive, just as in Yiddish, German, Japanese, and Korean. The sentences below show that Swedish object shift past the subject shows the same lack of reconstruction when the subject is not a potential binder for the scrambled pronoun.

(14) a. Igår gav Johans$_i$ syster honom$_i$ en bild.
Yesterday gave John$_i$'s sister him$_i$ a picture.
"Yesterday, John's sister gave him a picture."

b. Igår gav honom $*_{i/j}$ Johans$_i$ syster en bild.
Yesterday gave him$_{*i/j}$ John$_i$'s sister a picture.
(Christer Platzack, p.c.)

I conclude from the data above (and in the references cited) that the reconstruction facts simply do not prove that there are two types of scrambling, A and A', contra Mahajan (1990), and following Lee (1993) and Frank, Lee & Rambow (1996). In other words, the binding facts are entirely compatible with the view that scrambling is a unitary phenomenon, whether the landing site for a scrambled element is to the left of the subject or not. The data further suggest that one or both of the two following possibilities are true and should be explored: that reconstruction for binding is not a reliable

diagnostic for the A/A' distinction, and/or that the A/A'-distinction itself is not sufficient to exhaustively categorize all of the observed movement phenomena.

I would cautiously suggest that (at least) the latter is the case, and that scrambling does not fall neatly into the A/A' categorization because it actually is neither: A- and A'-movement are two types of movement distinguished by their targeting two different types of specifier positions, but scrambling is adjunction from an internal source (i.e. movement resulting in adjunction), and does not target specifier positions at all. Instead, scrambling, like all adjunction, always extends a maximal projection (in bare phrase structure terms this means a projection which already contains a specifier or one which does not require a specifier). The reason that scrambling is a challenge for the A/A'-distinction is merely because adjunction from an internal source is rare; adjunction is generally of modifiers, when they merge initially. Lee (1993) came very close to this conclusion in suggesting that scrambling was case-driven A-movement which shows reconstruction effects in certain cases. In fact, Lee cited Chomsky's (1993) three-way movement distinction which recognized adjunction as potentially different from A- and A'-movement, the distinction which I adopt here, but then proceeded to try and fold scrambling back into the A/A'-movement typology. This was a laudable attempt to simplify the theory, but in making this attempt, Lee was forced to make scrambling an odd type of A-movement chiefly because of two properties: it results in adjunction, and it shows some reconstruction for binding. Furthermore, as Lee (1993) admits, this stance is only tenable if all scrambled elements are required to receive Case, which is at least a nonstandard assumption where scrambled PPs and CPs are concerned. I would rather

bite the bullet and accept that the movement typology actually does include adjunction as a third type, and that it is the freedom inherent in adjunction (i.e. there are many possible adjunction sites in the phrase structure) which makes the Conservation of C-Command a necessary component of the grammar when adjunction is the result of movement, as it is with scrambling.

Lebeaux (1995) gives a possible answer which would serve to maintain the A'/A-scrambling distinction, at least with regard to some of the data: Lebeaux (1995) argues that Principle C applies at all stages of the derivation, and so a lack of reconstruction for the purposes of Principle C would be expected even if scrambling really were true A'-movement.[94] First of all, this is only a possible argument with regard to the Principle C data above; the argument from Principle A in (10) and (11) remains untouched. But as it turns out, Lebeaux turns out to be empirically wrong for a number of other reasons. The first among these is examples (1) and (3)-(5) above: if Principle C truly applies at all stages of the derivation, then the reflexive object pronouns which have been scrambled past the subject in these examples should bind the subject and produce ungrammaticality. They do not, so Principle C cannot apply on the surface in these examples.

Secondly, topicalization in English shows that Lebeaux's conclusion is wrong even for canonical cases of A'-movement. In sentences like the one below, a reflexive object which is coindexed with the subject has been topicalized, and the result is grammatical. It is clear that the object reconstructs not only for the purposes of Principle

[94] Thank you to Julie Legate for pointing this issue out to me.

A, but also for Principle C: if Principle C applied on the surface, the reflexive would also bind the R-expression in the subject and the sentence should be ungrammatical.

(15) Himself$_i$, John$_i$ likes.

While the sentence above is generally considered grammatical, English topicalization can be a bit awkward with a non-pronominal subject for independent, purely prosodic reasons (see Speyer 2005, 2008 for extensive data and discussion bearing on this issue). In order to cement the point (15) is intended to make, consider the sentences below as well, in which there is no confounding prosodic awkwardness.[95]

(16) He$_i$ tends to be skeptical of most faculty, but HIMSELF$_i$ my professor$_i$ actually believes.

(17) Mary told me that John$_i$ is skeptical about most people's intuitions, but himself$_i$ she's pretty sure John$_i$ usually believes.

(18) Here's the problem with John$_i$: while he$_i$ doesn't like people very much, HIMSELF$_i$ John$_i$ really HATES.

It is plain that in each of the examples above the reflexive pronoun c-commands a coindexed R-expression on the surface, which would result in a Principle C violation if the topicalized object did not reconstruct in its underlying position. Clearly

[95] In short, Speyer's work shows that the prosodic ill-formedness of sentences like (15) is due to a stress clash between the topicalized constituent and the non-pronominal subject. In (16), this issue is avoided by contrastively focusing the topicalized constituent, which simultaneously causes the subject to be deaccented, and the possessive pronoun *my* puts further distance between the heavier prosodic constituents. In (17), cyclic topicaliztion avoids this problem by placing the topicalized constituent next to a pronominal subject, even though it is bound by the R-expression *John* further on in the clause. Finally in (18), the double contrastive focus on *himself* and *hates* serves to deaccent *John*, much as in (16).

reconstruction does apply in these cases, saving the potential Principle C problem. Principle C cannot apply at all levels of the derivation, contra Lebeaux (1995).

8.3 The Reverse Webelhuth's Paradox

"Webelhuth's Paradox", so-called after Webelhuth (1989), is the observation that scrambling in German to landing sites below (to the right of) the subject appears to show properties of both A- and A'-movement simultaneously. On the basis of examples such as (19) below, Webelhuth proposed that scrambling constitutes a third type of movement in the languages that show it, movement to mixed A/A'-positions. In the example below, scrambling the quantified expression, *jeden*, creates a new binding relation and allows it to bind the possessive *seinem*, while simultaneously licensing an apparent parasitic gap in the position marked *e*.

(19) …daß Maria jeden$_i$ ohne *e* anzuschauen seinem$_i$ Nachbarn t$_i$
that Maria everyone$_i$ without to-look-at his$_i$ neighbor

vorgestellt hat
introduced has

"…that Maria introduced everyone to his neighbor without looking at him"

(Webelhuth 1989: 410)

Thus, scrambling below the subject does not show reconstruction for quantifier binding as in canonical A-movement (nor does it reconstruct for weak crossover; see also Lee

1994 for Korean, Lee & Santorini 1994 and Frank, Lee, & Rambow 1996 for German and Korean, Mahajan 1990 for Hindi), but paradoxically it appears to license parasitic gaps, which are generally taken to be licensed only by A'-movement since Engdahl (1983). Webelhuth (1989) uses this observation to argue that scrambling targets mixed A/A'-positions. If the interpretation of this effect is as clear as Webelhuth argues it is, then it challenges the analysis I argue for here, in which scrambling is not A'-movement (mixed or not), and in which there are therefore no mixed A/A' positions. I argue below that the German data should not be interpreted as Webelhuth suggests.

In Swedish, we can observe precisely the reverse effect from the one Webelhuth reports for German. Unlike Webelhuth's Paradox, this "Reverse Webelhuth's Paradox" can only be seen in examples of scrambling to a high position in Swedish, with the scrambled object landing to the left of the subject. As we have seen above, such "long-object-shifted" objects reconstruct both for the purposes of reflexive binding and for the purposes of (saving violations of) Principle C, provided that the subject is a potential binder for the scrambled object (as in German, Yiddish, and Japanese; see examples above in this chapter). Strikingly, the examples below show that the scrambling which results in reconstruction cannot license parasitic gaps.[96] Compare the first example in (20) below to the ungrammatical example which follows it. The contrast shows that while the first reflexive pronoun must reconstruct, the second is nevertheless obligatory and cannot be replaced with a parasitic gap.

[96] Note that Holbmerg & Platzack (1995) did show that object shift below the subject position does not license parasitic gaps (see also the discussion in Thráinsson 2001), but they did not specifically consider Swedish object shift to a position higher than the subject.

(20) a. I går diskvalificerade sig_i $Johan_i$ t_i utan att ändra sig_i
Yesterday disqualified $SELF_i$ $John_i$ without to change $SELF_i$
"Yesterday, John disqualified himself without changing his mind."

b. * I går diskvalificerade sig_i $Johan_i$ t_i utan att ändra <pg>
(example from Paul Kiparsky, p.c.)

The following example shows the same illicit configuration with another predicate:[97]

(21) * I går lade sig Johan utan att raka <pg>
Yesterday laid **SELF** John without to shave
Intended: "Yesterday John went to bed without shaving himself."

While the data from Swedish behaves exactly like the data from German in terms of the reconstruction facts, the judgments are exactly opposite the ones from German for parasitic gaps. Given the contradictory data about parasitic gaps, there are three possibilities: first, it is possible that Swedish object shift and German scrambling are not the same type of movement, as some have argued, and German scrambling really is A'-movment; or secondly, it might be that one of the two languages does not have true parasitic gaps.

But at this point, one can see that the evidence lines up neatly in a single direction. If we take scrambling past the subject to be A'-movement on the basis of the reconstruction facts, then we have created more problems than we have solved. First, there is no explanation for the lack of reconstruction when the subject is not a potential binder for the scrambled object, as shown by the Principle C violations in (6)-(9), (11)-(14) above in Korean, German, Japanese, and Swedish, as well as the ability of

[97] Judgments for these three sentences are from Paul Kiparsky and Christer Platzack, p.c.

scrambling to bind the reflexive possessive inside the subject shown in the Japanese sentences in (10)-(11). Lee & Santorini (1994), as well as Williams (2003: 129), make the wrong prediction in these cases by taking scrambling past the *subject* to be the crucial predictor of reconstruction, rather than scrambling past a *potential subject binder*, as we have seen above and in Lee (1993) and Frank, Lee, & Rambow (1996). And as I have noted above, it is easy to show that Principle C violations can be rescued by reconstruction in undisputed cases of A'-movement, contra one of the conclusions of Lebeaux (1995). It is clear from the data I have presented and cited that reconstruction in scrambling is dependent not on the movement, but rather on the status of the subject as a potential binder when the movement crosses the subject, as argued in Lee (1993).

Secondly, it has been accepted fact since Engdahl (1983) that A'-movement licenses parasitic gaps[98] (in addition to the empirical generalization, this fact is predicted by analyses of parasitic gaps involving null operator movement in the tradition of Chomsky 1986: 53). In light of this, maintaining that scrambling past the subject is A'-movement would constitute a single exception to this generalization in Swedish, and such an analysis should give the analyst pause. Note that the proposal in Webelhuth (1989), that scrambling targets mixed A/A'-positions, cannot solve the question of how Swedish scrambled reflexives can simultaneously reconstruct and fail to license parasitic gaps. Whereas Webelhuth's Paradox is a case of a position seemingly having too many properties, the Reverse Webelhuth's Paradox is a case of a position seemingly having too few properties, so a mixed A/A'-position in Webelhuth's sense (a conjunction of A/A'-

[98] To the skeptic I would remark that this is clear for English and Swedish, at the very least.

properties) continues to make the wrong prediction for this case. In light of the Swedish data, Webelhuth's proposal of a mixed position complicates the theory by positing an entirely new movement type without any gain in empirical coverage.

One might argue that Scandinavian "object shift" is simply different from German-style scrambling, but on top of giving up on the results that the GHC/Conservation of C-Command represents, such an analysis would ignore the observation that the binding facts for "long object shift" in Swedish are identical to those for scrambling to the same structural position in German, Korean, and Japanese. On the other hand, accepting the fact that reconstruction does not constitute a perfect diagnostic for A'-movement allows scrambling to be unified in a single analysis (the GHC), does not disturb the results of Engdahl (1983) and subsequent researchers on parasitic gaps, and also allows the observations concerning reconstruction down A-movement chains in Lebeaux (1995) to stand. For the relatively small cost of abandoning reconstruction as a hard-and-fast diagnostic for A'-movement, which Lebeaux (1995) shows we have to accept anyway, we gain a better understanding of parasitic gaps and a unified analysis of scrambling.

Now the only clear challenge to analyzing scrambling as a unitary phenomenon is Webelhuth's original paradox, the licensing of parasitic gaps in German scrambling. The status of apparent parasitic gaps in German is not a problem that I can possibly attempt to solve in the present work. However, I will simply note here that there is no consensus in the literature on whether true parasitic gaps actually exist in German and Dutch, and in the absence of a clear result on the issue, I will assume that the apparent parasitic gap

constructions in German and Dutch are not currently a reason to reject an analysis of scrambling as a unitary operation. The problem is particularly difficult since, in addition to the general lack of preposition-stranding in German, which simply removes a possible environment for the investigation of parasitic gaps, the distribution of apparent parasitic gaps in German is generally so restricted that it is difficult to replicate the environments described in Engdahl (1983) and subsequent work on English and Swedish parasitic gaps (see Culicover 2001 for an overview of this research area).

In fact, Kathol (2001) considers the highly restricted distribution of apparent German parasitic gap constructions itself to be an argument against analyzing them as parallel to English and Swedish parasitic gaps. While I do not necessarily endorse the details of Kathol's final analysis, he shows that apparent parasitic gaps in German are constrained in ways that are not configurational in nature (e.g. lexical effects), and plausibly represent a different phenomenon from true parasitic gaps. Kathol (2001: 323) notes that this construction is limited to the *zu*-infinitival complements of a few specific prepositions, *ohne* ("without", (*an*)*statt* ("instead"), and *um* ("in order to"), an unexpected fact if the gaps in question are true parasitic gaps licensed by A'-movement in the relevant configuration. Some constraints on apparent German parasitic gaps are non-lexical in nature, but do not have an obvious configurational explanation either, as in the following contrast (Kathol 2001: 326, originally due to Mahajan 1990: 56):

(22) a. ??Peter hat jeden Gast$_i$ seinem$_i$ Nachbarn ohne <pg> anzuschauen
Peter has each guest$_i$ his$_i$ neighbor without to-look-at

vorgestellt.

introduced

b. Peter hat jeden Gast$_i$ ohne <pg> anzuschauen seinem Nachbarn
Peter has each guest$_i$ without to-look-at his$_i$ neighbor

vorgestellt.
introduced
"Peter introduced each guest to his neighbor without looking at him (= the guest)."

Similar observations were made about Dutch by Huygbregts & van Riemsdijk (1985), and they also conclude that Dutch apparent parasitic gaps are actually instances of a different construction than the English and Swedish versions.

Bayer & Kornfilt (1994: 24-25) also note a contrast that is surprising under the view that scrambling licenses true parasitic gaps in German. The following examples show contexts in which scrambling of a definite (demonstrative) DP cannot license an apparent parasitic gap, but scrambling of an object pronoun can.

(23) a. Man hat ihn$_i$ ohne <pg$_i$> verwarnt zu haben t$_i$ ins Gefängnis
One has him without warned to have in-the prison

gesteckt.
put
"One has put him into jail without having warned (him)."

b. ?* Man hat diesen Mann$_i$ ohne <pg$_i$> verwarnt zu haben t$_i$ ins
One has this man without warned to have in-the

Gefängnis gesteckt.
prison put
Intended: "One has put this man into jail without having warned (him)."

(24) a. Da hat ihn$_i$ der Polizist ohne <pg$_i$> verwarnt zu haben t$_i$ ins

There has him the policeman without warned to have in-the

Gefängnis gesteckt.
prison put
"The policeman has put him into jail without having warned (him)."

b. * Da hat diesen Mann$_i$ der Polizist ohne <pg$_i$> verwarnt zu
There has this man the policeman without warned to

haben t_i ins Gefängnis gesteckt.
have in-the prison put

On the basis of these data Bayer & Kornfilt (1994: 24-25) suggest that the apparent parasitic gap construction is not a reliable diagnostic for A'-movement in German, since there is no *a priori* reason to expect pronoun scrambling to behave differently from DP-scrambling in structural terms alone (though it is unclear whether or not they believe these to be true parasitic gaps). Haverkort (1993) addresses this issue, arguing that "pronoun scrambling" is in fact a type of cliticization, and cliticization licenses parasitic gaps in Germanic while scrambling does not. This argument is completely untenable on empirical and conceptual grounds. First, Scandinavian weak pronoun object shift does not license parasitic gaps, neither when the pronoun is moved past the subject (as in the Swedish examples above), nor when it lands in a lower position (Holmberg 1986: 173, Holmberg & Platzack 1995: 146), and Scandinavian weak pronouns are frequently as phonologically enclitic (or even more reduced) than their West Germanic counterparts (cf. Hellan & Platzack 1995). Secondly, Haverkort's treatment is undesirable on conceptual grounds: he defines Germanic clitics as maximal projections adjoined to a maximal projection (1993: 131), which is simply a renaming of pronoun scrambling

(under the analysis pursued here, in Lee 1993 (and references therein), and Cardinaletti & Starke 1999 who carefully distinguish pronoun scrambling from true cliticization).

In summary, while Webelhuth's Paradox shows scrambling to have an apparently dual A/A' nature, the Reverse Webelhuth's Paradox shows that the picture in reality is more complex, and both the parasitic gap diagnostic and the reconstruction diagnostic for A'-movement should be reconsidered. I have shown that on reflection, the reconstruction effect is real, but not necessarily a proof that certain types of scrambling are actually A'-movement. The evidence from parasitic gaps in Swedish, on the other hand, is entirely consistent with scrambling (whether past the subject or to a lower landing site) being some operation other than A'-movement, and I take this operation to be internal adjunction. Note that the data is most consistent with respect both to reconstruction and parasitic gaps in the language that uncontroversially allows parasitic gaps, while the two sets of data only appear to be contradictory in German and Dutch, where the evidence for the existence of parasitic gaps is widely debated. While this is not direct evidence that parasitic gaps do not occur in German and Dutch, I would suggest that when a case of controversial and difficult to interpret findings conflicts with a clear pattern, the best research policy is to pursue the clear pattern until better data about the other cases becomes available. In this spirit, I will assume that the parasitic gap data does not necessarily cut against the theory I have pursued here, and may in fact be in favor of it in the case of Swedish.

8.4 Summary

In this chapter I have suggested that the distinction frequently made between A- and A'-scrambling is both unnecessary and based on very shaky evidence. While both the reconstruction facts and parasitic gap facts are difficult to interpret, neither is an obstacle to a unified analysis of scrambling as an operation of adjunction from within. Instead, the evidence from reconstruction ultimately becomes evidence for a further unification in the theory, rather than a division: the reconstruction differences for local (i.e. below subject) vs. non-local (above subject) scrambling reappear in Swedish object shift, lending more weight to the idea that these phenomena should be unified. The comparison between Swedish and Korean, Japanese, and German also brought to light a hitherto unnoticed phenomenon, the Reverse Webelhuth's Paradox, which definitively shows that scrambling/object shift past the subject in Swedish (and probably more generally) is not A'-movement.

Chapter 9

A Prima Facie Exception to the GHC: Middle English Clitic Pronouns

9.1 Introduction: The Problem

English begins with a scrambling system like modern German's, in the early Old English of *Beowulf* (Pintzuk 1996), and then changes to show the same scrambling patterns that are found in modern Yiddish and Swedish, in different stages of its history (as I discussed above). Additionally, English showed one other system that is otherwise unattested in Germanic; some pronoun objects appear to the left of the finite verb in Tense, as in (1) and (2) below, in apparent violation of the Generalized Holmberg Constraint / Conservation of C-Command. This chapter presents a new analysis of this "pre-Tense" pronoun system, discussing the syntax of these pronouns in Middle English, for which there is a great deal of information about syntactic variation across dialect boundaries. If this phenomenon is a type of phrasal scrambling, then it constitutes an obvious, albeit rare, counterexample to the GHC/Conservation of C-Command (which I will continue to refer to with the acronym GHC in this chapter, even though its

formulation has changed in the "Conservation of C-Command"). If the GHC is to be maintained as a universal, then pre-Tense pronouns must arise by some other process.

I argue that the GHC in fact makes precisely the correct prediction in this case: rather than assuming the rather *ad hoc* analysis that pre-Tense pronouns are a typologically aberrant form of scrambling, the GHC forces a more sensible analysis in which the pre-Tense pronouns are an instance of a known phenomenon that is distinct from phrasal scrambling. In short, if pre-Tense pronouns cannot be derived by phrasal movement (scrambling), then they must be derived by head-movement, or cliticization (in the sense of Kayne 1991). Unlike the usual Germanic scrambling of objects, pre-Tense pronouns are true clitic pronouns, which are head-adjoined to Tense. This Tense-clitic system appeared some time during the Old English period, and disappeared gradually during the Middle English (ME) period. During the course of this change, weak object pronouns with leftward scrambling remained in the language, but occurred at different frequencies depending on the proportion of Tense-clitics that were in use at a given time.

9.2 Three positions for pronoun objects (where only two are expected)

Kroch & Taylor (2000a) established that there was considerable variation in the structure of vPs in ME , which could be underlyingly left-headed or right-headed, and that the base order of vPs was frequently masked in clauses with pronominal objects, by

the frequent occurrence of a construction in which the pronoun appeared in a position immediately to the left of the finite verb or auxiliary ("pre-Tense" position); this is shown in the examples below.

(1) and he **hit** wille do bleþeliche.
and he it will do blithely.
" and he'll gladly do it."
(*The Kentish Sermons*; CMKENTSE,219.149 in the PPCME2)

(2) þa feala þing **hire** byð wiðtogene,
those many things her-DAT are taken away.
"Many things were taken away from her"
(*The Kentish Homilies*; CMKENTHO,138.116)

The pre-Tense position for object pronouns is well-known from studies of Old English (cf. van Kemenade 1987, Pintzuk 1991, and Pintzuk 1996), and in the Middle English period, it is still the dominant position for weak pronouns in the more conservative Middle English texts, such as *The Kentish Sermons* and *The Kentish Homilies*. When one considers texts outside of Kent, it becomes clear that the early ME period (c.1100-1350) saw robust variation in the placement of object pronouns, which could be found in two additional positions as well: "post-Tense" and "post-Verb," which are illustrated below.

(3) & he **hit** wule ȝelde **ȝe** as his treowe feire. wið halewi of heouene.
and he it will yield you as his true companion with balm of Heaven.
"And he will grant it to you along with the balm of Heaven, as his true companion."
(*The Ancrene Riwle*; CMANCRIW-1,II.91.1099)

(4) I, forr þe lufe off Godd & forr þe mede off heffne, **Hemm** hafe **itt**
I, for the love of God and for the reward of heaven, them have it

inntill Ennglissh wennd Forr þe*33*re sawle 272olk.
into English turned for their soul need.

"I, for the love of God and for the reward of heaven, have translated it into English for them on account of their souls' need."
(*The Ormulum*; CMORM,DED.L143.36)

The GHC suggests a specific hypothesis concerning the structures underlying each of the three possible positions for weak object pronouns that are found in Middle English. If scrambling cannot take place across an intervening head, then post-Tense pronouns must originate to the left of the verb (in an OV *v*P) and post-Verb pronouns originate to the right of the verb (in a VO *v*P). What then is to be made of pre-Tense pronouns, since they have certainly moved from their theta-positions to cross Tense, a c-commanding head with a merged finite verb? If the Generalized Holmberg Constraint is at all valid, then these pronouns cannot represent scrambling in the sense I have described above, i.e., as a general leftward movement to a sequence of phrasally-adjoined landing sites. There is only one analytic option left if the GHC is to be maintained: if the pre-Tense pronouns are not phrasal, then they must be heads. I suggest that pre-Tense pronouns are head-adjoined clitics on Tense ("true clitics" in Cardinaletti & Starke 1999), and so have landed in their surface positions not by scrambling, but via head-movement in the same way that Romance clitics do (cf. Kayne 1991).

The three positions for weak object pronouns in English are not stable diachronically, and ultimately, both the pre-Tense and post-Tense pronouns disappear from the language. While the change is in progress in the early ME period (pre-c.1350),

different ME dialects show different stages of this change, with the Kentish and Southeastern dialect areas showing the earliest stage of the change. (The basic diachronic pattern is shown in the chart below, corresponding to Tables 9 and 10, with the older and more conservative texts appearing on the left, and the most innovative ones on the right. This represents only a small portion of the gathered data, of course.) Kentish and the Southeastern texts are indepently known to represent earlier stages of the language, compared with the more northerly and westerly dialects; their conservative nature can be observed outside of the syntax in a number of phonological changes and the loss of morphological case-marking (Allen 1995). Moving across England northward and westward from Kent, from the Southeast Midlands to the Northeast Midlands and then West Midlands, and forward in time, the use of the pre-Tense position for object pronouns gradually recedes in favor of first, the post-Tense position, and ultimately, the post-V position.

By the year 1350 the loss of pre-Tense clitics had gone to completion, leaving in its wake Late ME dialects with head-initial *v*Ps and object-shift of the modern mainland Scandinavian type, as I discussed at length in Chapter 4 above. Just as the GHC makes a specific prediction about the structure of pre-Tense clitics, it also forces a particular analysis of this change in the position of weak pronouns. Since post-Tense and post-Verb pronouns are in the predicted positions for weak pronoun objects in underlyingly OV and VO *v*Ps, the change must be seen as a competition between two grammars, not three. The two systems are the Tense-adjoined (head-adjoined) clitic grammar, and the usual Germanic weak pronoun grammar, which entails leftward scrambling of weak

pronouns as phrases and bounded by the position of the finite verb in Tense, according to the GHC. The Tense-adjoined clitic grammar is lost during the Middle English period, and it just so happens that the OV to VO change in the structure of the *v*P is progressing at the same, making it appear as if pronouns in Middle English are of three types. In fact, there are only two types of (non-strong) pronoun, clitic or weak, but the latter type will appear either in a post-Tense or a post-Verb position depending on whether the particular clause it originates in is underlyingly OV or VO.

9.3 The Interaction between the OV-to-VO change and pronoun position

The distribution of weak pronouns and clitics in different ME texts is consistent, in detail, with this particular view of the change in pronoun syntax. In particular, pre-Tense clitics occur at high frequencies even in texts that are predominantly VO underlyingly, where the frequencies of OV and VO are independently estimated by observing the placement of full DP objects. This shows that pre-Tense clitics are not grammatically tied to the OV system. Post-Tense pronouns, on the other hand, are dependent on the OV grammar, and as expected, they eventually disappear as OV is lost. However, there is a slight rise in post-Tense pronoun frequency in the center of the graph below, which is unexpected given that true OV structure declined steadily across the texts shown in the graph (see Kroch & Taylor 2000a). This suggests that the use of pre-Tense clitics bleeds both the frequencies of post-Tense and post-V pronouns, which we would

expect if pre-Tense clitics are in competition with all weak pronouns (OV or VO). Then as pre-Tense pronouns decline, there is an expected rise in the frequencies of both post-Tense and post-V pronouns. Examples such as (3) and (4) above certainly suggest this view of things, as it is obviously possible for pre-Tense pronouns to co-occur with either post-Tense or post-V pronouns. The corpus contains three more examples showing the possibility of pre-Tense pronouns in unambiguously VO clauses, and I have repeated these below.

(5) & icc **itt** hafe forðedd **te**, Acc all þurrh Cristess hellpe;
And I it have defeated thee, yet all through Christ's help
"I have defeated it for you, yet through the help of Christ"
(*The Ormulum*; CMORM,DED.L23.8)

(6) þt ich **hit** habbe itald þ**e**.
that I it have told thee
"that I have told you it."
(*The Ancrene Riwle*; CMANCRIW-1,II.76.893)

(7) Whær icc **me** mu3he findenn **himm** To lakenn himm & lutenn
where I me might find him to workship him and prostrate
"where I might find him for myself, to worship him and prostrate myself to him"
(*The Ormulum*; CMORM,I,222.1834)

Even without any degree of quantitative sophistication, the slight rise in the frequency of post-Tense pronouns in the middle of the decline of pre-Tense pronouns has an immediate explanation under this hypothesis. If pre-Tense bleeds both post-Tense and post-V, then the loss of the pre-Tense position should cause the use of the other two positions to rise, in proportion to the amount of underlyingly OV and VO *v*Ps a given text

contains. This explains the difference between the *Kentish Sermons* and the *Kentish Homilies*, as shown on the left side of Figure 5 below and the corresponding table below it. If OV to VO and pre-Tense clitics are not grammatically connected, then a change in one does not necessarily need to affect the other. In this case, the Kentish dialect has managed to remain conservative over a 150 year period (roughly; i.e., from the *Homilies* to the *Sermons*) in requiring most pronouns to occur as pre-Tense clitics, while it has nevertheless undergone the shift from OV to VO. Thus, even though most pronouns presumably originate in VO *v*Ps in the *Kentish Sermons*, both post-Tense and post-V positions are bled by the pre-Tense pronouns in both the *Sermons* and the *Homilies*, and so the two texts appear similar in the syntax of pronominal objects only.

It is an open question as to why one change should have spread to the author of the *Sermons* while the other did not, but given that it obviously did so, the only tenable hypothesis concerning the loss of OV and the loss of pre-Tense pronouns is that they are not reflexes of a single underlying syntactic change, but rather reflect a situation in where two grammatical components are changing. *Vices and Virtues* contrasts with the *Kentish Homilies* in the same way as the *Kentish Sermons*, with the benefit that the overall N is much higher for *Vices and Virtues* than it is for the *Kentish Sermons*. According to the counts of full DPs, *Vices* is 61.7%-77.8% VO, depending on clause type, but this does not prevent *Vices* from realizing 71.9%-74.4% of its pronominal objects as pre-Tense clitics. *The Trinity Homilies*, on the other hand, shows more OV with full DPs than *Vices* does, 25.35% and 55.56% for matrix and subordinate clauses, respectively, but contains a lower proportion of pre-Tense pronouns than *Vices*: *Trinity* has 40.63% and 58% pre-

Tense pronouns, compared with *Vices*' 71.9% and 74.4%. These four texts, *The Trinity Homilies*, Vices *and Virtues*, *The Kentish Homilies*, and *The Kentish Sermons* are sufficient in themselves to show the independence of the OV to VO change and the loss of pre-Tense clitics: they all show nearly the same distribution of pronominal objects, even though they all have different rates of OV with DP objects (comparing the frequencies found in subordinate clauses). The data from *Vices and Virtues* and the *Kentish Sermons* also makes it unlikely that post-Tense pronouns could reflect scrambling across the verb from an underlying VO position. Their low rates of post-Tense pronouns accurately reflect the low rates of OV in these texts, as measured by full DPs, minus those pre-Tense pronouns which originated in an OV position. The comparison between the rate of OV as estimated by DP position and the rate of pre-Tense pronouns may be seen in the chart below.

Figure 5. Pre-Tense pronouns vs. OV with full DPs for four texts (subordinate clauses)

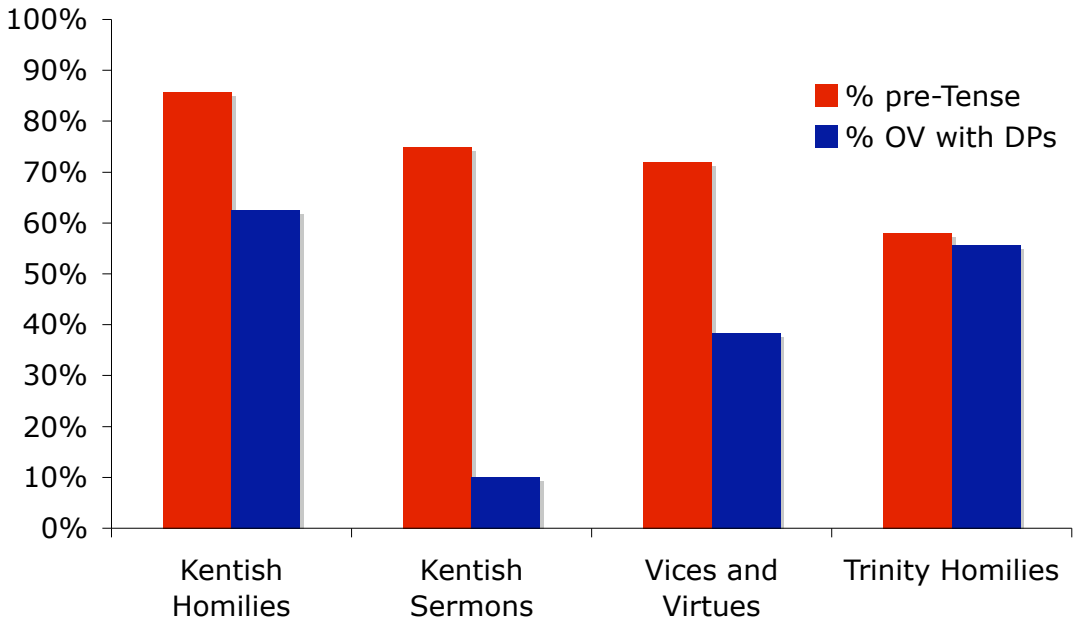


Table 8.

Southeastern Texts	% pre-Tense pronoun objects (see tables below)	aux > DP > V	aux > V > DP	% OV with DP objects
Kentish Homilies	**85.71%**	5	3	**62.50%**
Kentish Sermons	**75.00%**	2	18	**10.00%**
Vices and Virtues	**71.95%**	36	58	**38.30%**
Trinity Homilies	**58.00%**	45	36	**55.56%**

It is possible to test this hypothesis more rigorously in the following way. Using loglinear models (see Agresti 2002 for a description and full discussion of inference for loglinear models), one can define models corresponding to the hypothesis that OV is statistically (and therefore grammatically) tied to the use of pre-Tense pronouns, and then check the plausibility of such a model against the observed distribution of data. As I mentioned above, it is already clear from the Late Middle English situation that pre-Tense pronouns cannot be a feature of the VO system. In order to statistically test whether pre-Tense pronouns are connected to OV vPs, I recoded the data for Early Middle English subordinate clauses in terms of 3 variables: text (T), position (P): "pre" or "post", and object is a DP or pronoun (D). In order to be able to directly compare the frequencies of DPs and pronouns, I have collapsed the pronoun data into the two categories "pre" and "post", for "preceding some verb" and "following some verb". Thus, for each text (each value of the "text" variable), the combination of P = "pre" and D = DP is the count of DPs in OV position, P = "post" and D = DP is the count of DPs in VO position, P = "pre" and D = pronoun is the count of pre-Tense pronouns, and the count of P = "post" and D = pronoun is the sum of post-Tense and post-Verb pronouns for that text. With the variables defined in this way, I was able to investigate the

relationship between the frequency of pre-Tense pronouns and the frequency of DPs in OV position by testing loglinear models that specify a certain relationship between the D and P variables.

The first model I tested allowed main effects for each variable, and then interactions between text and D, and between text and P, but no interaction between P and D; this model estimates the counts for the different combinations of variables under the assumption of conditional independence between position and whether the object is a DP or a pronoun, given a particular text. (In the notation commonly used for describing loglinear models, this model is [TD][TP]). In other words, this model should fit the observed data well if for any given text, pronominal objects and DP objects are equally likely to appear in their respective "pre" positions; i.e., this model tests the strong hypothesis that the pre-Tense position is simply a type of OV.

Somewhat unsurprisingly, the model of conditional independence between position and "DP or pronoun" does not fit very well. The Pearson and G^2 tests of fit are above 49 on 8 degrees of freedom for this model: $p < 5 \times 10^{-8}$. Clearly, treating pre-Tense for pronouns and OV for DPs as the same thing statistically does not yield a distribution that is very close to the observed data. However, there is one more model corresponding to an hypothesis under which OV and pre-Tense are related in some way. This model includes the same terms as the previous one, but with one added term: a term for an interaction between P and D. This is known as the homogenous association model, or the "no 3-way interaction" model in the statistical literature, and it has the property that the odds ratios between each pair of variables is constrained to be constant at each

level of the third. For the problem at hand, the important prediction of this model is that it allows the odds of an object being in "pre" or "post" position to be different depending on whether the object is pronominal or a DP; this was not the case for the previous model, which assumed conditional independence between the P and D variables (i.e. an odds ratio of 1 for each text). However, it constrains these odds to be the same for each text. In terms of linguistic theory, this model tests the hypothesis that OV and pre-Tense are not the same thing, but rather that pre-Tense for pronouns and OV for DPs are related in a systematic way across all of the texts. If pre-Tense position is a grammatical option that is only available in the context of an OV *v*P (even if it is distinct from underlying OV), then the level of pre-Tense pronouns should vary from text to text along with the level of OV with DPs in each text, but it should not vary from text to text in wildly different ways from the frequency of OV with DPs. The homogenous association model tests whether the observed data is plausibly derived from this type of relationship between the text, position, and "DP or pronoun" variables.

It is not. The Pearson and G^2 statistics for the fit of the homogenous association model to the subordinate clause data are 19.5 and 19.9 on 7 degrees of freedom, $p < .01$. Clearly not a very likely model, then. The badness of fit results from the fact that the frequency of OV with DPs and the frequency of pre-Tense pronouns are quite far from changing together in lockstep as you move from text to text.

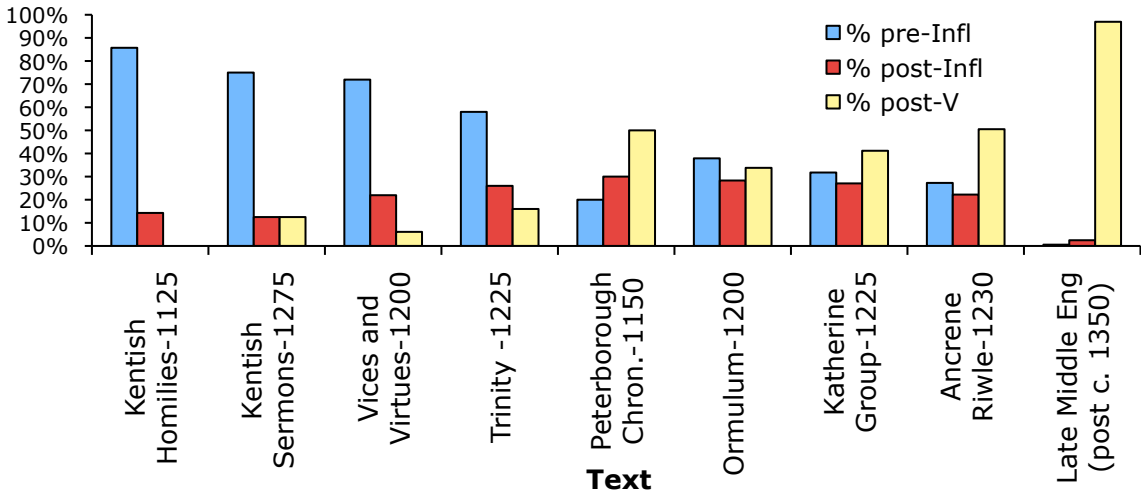
Table 9.
Pronoun placement in Early Middle English Clauses with Auxiliaries (Subordinate Clauses)

Kentish:	**pre-Tense**	**post-Tense**	**post-V**	**Total**	**% pre-Tense**	**% post-Tense**	**% post-V**
Kentish Homilies-1150/1125	6	1	0	7	85.71%	14.29%	0.00%
Kentish Sermons-1275	6	1	1	8	75.00%	12.50%	12.50%
Southeast Midlands:							
Vices and Virtues-1225/1200	59	18	5	82	71.95%	21.95%	6.10%
Trinity -1225	29	13	8	50	58.00%	26.00%	16.00%
Northeast Midlands:							
Peterborough Chron.-1150	2	3	5	10	20.00%	30.00%	50.00%
Ormulum-1200	83	62	74	219	37.90%	28.31%	33.79%
West Midlands:							
Katherine Group-1225	27	23	35	85	31.76%	27.06%	41.18%
Ancrene Riwle-1230	27	22	50	99	27.27%	22.22%	50.51%

Table 10.
Pronoun placement in Late Middle English (post-1350) Clauses with Auxiliaries, (Subordinate Clauses)

Late ME	**pre-Tense**	**post-Tense**	**post-V**	**Total**	**% pre-Tense**	**% post-Tense**	**% post-V**
Southern Texts	1	7	75	83	1.21%	8.43%	90.36%
East Midlands Texts	4	5	735	744	0.54%	0.67%	98.79%
West Midlands Texts	3	13	319	335	0.89%	3.88%	95.22%
Northern Texts	3	6	159	168	1.79%	3.57%	94.64%

Figure 6.
Pronoun placement in Early Middle English Clauses with Auxiliaries (Subordinate Clauses)



9.4 Conclusions

The statistical effects reported above are expected if the availability pre-Tense position is grammatically independent of whether a speaker's *v*Ps are underlyingly OV or VO, which is the case if pre-Tense position reflects a type of cliticization rather than weak pronoun scrambling. This would make texts like *Vices and Virtues* and *The Kentish Sermons* similar to modern French or Italian, which are categorically VO and have pre-Tense clitics. More importantly for the purpose of this dissertatin, the analysis favored by the quantitative data also has the consequence that pre-Tense pronouns are not a type of scrambling that violates the GHC, but rather represent a different type of movement, head movement to Tense.

The change in pronominal object syntax during the ME period can then be understood as reflecting two underlying changes in a cascade. Pre-Tense clitics and VO

vPs must have come into the English speech community at some point during the Old English period, and both were lost by the end of the Middle English period. Pre-Tense clitics are used in the place of weak pronouns when they occur, lowering the overall frequency of weak pronouns, regardless of whether these originate in right-headed or left-headed vPs. The weak pronouns that remain can occur to the right of the auxiliary (post-Tense) in an OV clause, whether they are scrambled or not, following the GHC. They can also occur to the right of the nonfinite verb just in case the clause is VO, in which case their movement is limited by the position of the nonfinite verb. Pre-Tense clitics begin as the preferred way to realize unstressed pronominal objects in the most archaic early Middle English. As pre-Tense clitics decline, but OV remains at an appreciable frequency, the post-Tense position asserts itself, as we can see in the contrast between the Kentish and the (North- and South-)east Midlands texts. At the same time, VO vPs are overtaking OV ones, and the frequencies of weak pronouns shift again, in favor of the VO post-V position; this is the contrast between the East Midlands texts and West Midlands texts, and ultimately between early ME, late ME, and Early Modern English.

The fact that pre-Tense pronouns also decrease over time then represents an independent change in English, and the diachronic pattern in pronoun distribution derives from how the two changes, OV to VO and the loss of Tense-clitics, overlap. However, without a clear and fine-grained theory of which types of scrambling are allowed by UG, it would not necessarily have been possible to disentangle the two changes. In this way, the fact that the GHC/Conservation of C-Command led to a more precise interpretation of the diachronic data counts as indirect, but important evidence in its favor.

Chapter 10

Conclusion: A Note on the Autonomy of Syntax

This dissertation accounts for the typology of scrambling found in a number of languages and during periods of language change, and also shows that the way in which scrambling is constrained provides insight into basic properties of phrase structure. In addition, the dissertation constitutes an extended argument for the autonomy of syntax: while prosodic and pragmatic considerations favor leftward scrambling in a number of contexts, a language's inventory of functional heads nevertheless puts a strict upper bound on whether scrambling can respond to these considerations. Thus, languages with more head-final clausal structure can use scrambling to signal information structure and accommodate the prosodic contours which accompany different information structural configurations to a much larger extent than head-initial languages can.

But this is precisely what is expected in a grammar with true modularity: the wide functionality of scrambling in languages like Korean or German does not mean that it must have wide functionality in a language like English, if independent syntactic

constraints exist in English which prevent scrambling from taking place in most contexts. In contexts where languages agree that scrambling can take place, it may serve the same purposes, but in other contexts, some more head-initial languages are simply handicapped. When this is the case in one area of syntax, these languages must compensate with other syntactic operations, or with purely prosodic ones. This is also a potential answer to the functionalist perspective which might question why UG would ever allow the seemingly baroque derivations of head-final languages, if the Kaynian approach is correct: the derivations may restrict the syntax in some areas, but they free up the syntactic apparatus for scrambling.

Bibliography

Abraham, Werner and Anko Wiegel. 1993. Reduktsionsformen und Kasussynkretismus bei deutscen und niederländischen Pronomina. In Werner Abraham and Josef Bayer eds., *Dialektsyntax*, Opladen: Westdeutscher Verlag.

Agresti, Alan. 2002. *Categorical Data Analysis 2nd Edition*. Hoboken: Wiley.

Allen, Cynthia. 1995. *Case Marking and Reanalysis: grammatical relations from Old to Early Modern English*. Oxford: Oxford University Press.

Anonymous. 1998 [1975]. *Living Sober*. New York: Alcoholics Anonymous Word Services, Inc.

Arnold, Jennifer E., Anthony Losongco, Thomas Wasow, and Ryan Ginstrom. 2000. The Effects of Structural Complexity and Discourse Status on Constituent Ordering. *Language* 76:1, 28-55.

Asudeh, Ash. 2000. Argument structure and animacy restrictions on anaphora. In Steve S. Chang, Lily Liaw and Josef Ruppenhofer, eds., *Proceedings of the Berkeley Linguistics Society 25*. Berkeley, CA: Berkeley Linguistics Society. 2–13.

Bade, Markus, Tanja Schmid, and Jana Häussler. 2009. Optionality in verb cluster formation. In Sam Featherston and Susanne Winkler eds., *The Fruits of Empirical Linguistics*. The Hague: Mouton de Gruyter. pp. 37-58.

Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: Chicago University Press.

Baltin, Mark. 1982. A Landing Site Theory of Movement Rules. *Linguistic Inquiry* 13, 1: 1-38.

Baltin, Mark. 2006. Extraposition. In Martin Everaert & Henk van Riemsdijk eds., T*he Blackwell Companion to Syntax, vol. II*. Oxford: Blackwell.

Bayer, Josef. 2006. Wh-in-Situ. In Martin Everaert & Henk van Riemsdijk eds., T*he Blackwell Companion to Syntax, vol. V*. Oxford: Blackwell.

Bayer, Josef and Jaklin Kornfilt. 1994. Against scrambling as an instance of Move-alpha. In Norbert Corver and Henk van Riemsdijk eds., *Studies on Scrambling: Movement and Non-movement Approaches to Free Word-Order Phenomena*. Berlin: Mouton de Gruyter.

Besten, Hans den. 1983. On the interaction of root transformations and lexical deletive rules. In Werner Abraham ed., *On the Formal Syntax of Westgermania*, p.47-131. Amsterdam: John Benjamins.

Besten, Hans den and Hans Broekhuis. 1992. Verb Projection Raising in het Nederlands. *Spektator* 21,1: 1-34.

Besten, Hans den and Corretje Moed-van Walraven. 1986. The Syntax of Verbs in Yiddish. In Hubert Haider and Martin Prinzhorn eds., *Verb Second Phenomena in Germanic Languages*. Dordrecht: Foris, p.111-125.

Bhatt, Rakesh M. 1999. *Verb Movement and the Syntax of Kashmiri*. Kluwer: Dordrecht.

Biberauer, Theresa. 2003a. *Verb second (V2) in Afrikaans: a minimalist investigation of*

word order variation. Ph.D. dissertation, University of Cambridge.

Biberauer, Theresa. 2003b. Reconsidering the EPP and Spec-TP in Germanic. In Astruc, Luisa & Marc Richards eds., *Cambridge Occasional Papers in Linguistics* (COPiL) 1: 100-120.

Biberauer, Theresa and Ian Roberts. 2005. Changing EPP-parameters in the history of English: accounting for variation and change. *English Language and Linguistics* 9, 1: 5 – 46.

Bies, Ann Elizabeth. 1996. *Syntax and Discourse Factors in Early New High German: Evidence for Verb-Final Word Order*. Master's Thesis, University of Pennsylvania.

Bobaljik, Jonathan David. 2005. Re: CycLin and the role of PF in Object Shift. *Theoretical Linguistics* 31: 111-125.

Borsley, Robert D., Maria-Luisa Rivero, and Janig Stephens. 1996. Long head movement in Breton. In Robert D. Borsley and Ian Roberts eds., *Syntax of the Celtic Languages*. Cambridge: Cambridge University Press.

Bresnan, Joan and Lioba Moshi. 1990. Object Asymetries in Comparative Bantu Syntax. *Linguistic Inquiry* 21, 2: 147-185.

Cardinaletti, Anna and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk Van Riemsdijk, ed., *Clitics in the Languages of Europe*. Berlin: Mouton de Gruyter.

Chomsky, Noam. 1970. Remarks on Nominalization. In Jacobs, R. and P. Rosenbaum eds., *Readings in English Transformational Grammar*, Waltham, MA: Blaisdell.

Chomsky, Noam. 1986. *Barriers*. Cambridge: MIT Press.

Chomsky, Noam and Howard Lasnik. 1993. The theory of principles and parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Venneman, eds., *Syntax: An international handbook of contemporary research*. Berlin: Walter de Gruyter.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In K. Hale and S. J. Keyser eds., *The View from Building 20*: *Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press. 1–52.

Chomsky, Noam. 1995. Bare Phrase Structure. In Gert Webelhuth ed., *Government and Binding Theory and the Minimalist Program*. Cambridge: Blackwell. 383–439.

Chomsky, Noam. 1998/2000. Minimalist Inquiries: The Framework. *MIT Occasional Papers in Linguistics* 15. Republished in 2000 in R. Martin, D. Michaels, & J. Uriagereka (eds.), *Step By Step: Essays In Syntax in Honor of Howard Lasnik*. Cambridge, Mass: MIT Press. 89–155.

Cognola, Federica. 2007. The loss of OV patterns in the German dialect spoken in the Fersina Valley. Paper presented at the ScanDiaSyn Grand Meeting, August 16-19, Lake Mývatn, Iceland.

Condoravdi, Cleo and Paul Kiparsky. 2002. Clitics and Clause Structure. *Journal of Greek Linguistics* 2,1: 1-39.

Condoravdi, Cleo and Paul Kiparsky. 2004. Clitics and Clause Structure: The Late Medieval Greek system. *Journal of Greek Linguistics* 5,1: 159-183.

Culicover, Peter W. 2001. Parasitic Gaps: A History. In Peter W. Culicover and Paul M. Postal eds., *Parasitic Gaps*. Cambridge: MIT Press.

Diesing, Molly. 1990. Verb Movement and the Subject Position in Yiddish. *Natural Language and Linguistic Theory* 8.1: 41-79.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Diesing, Molly. 1997. Yiddish VP Structure and the Typology of Object Movement." *Natural Language and Linguistic Theory* 15.2: 369-427.

den Dikken, Marcel. 1995. *Particles: on the syntax of verb-particle, triadic, and causative constructions*. Oxford: OUP.

É. Kiss, Katalin. 2002. *The syntax of Hungarian*. Cambridge: Cambridge University Press.

Embick, David and Rolf Noyer. 2001. Movement Operations after Syntax. *Linguistic Inquiry* 32, 4: 555-595.

Engdahl, Elisabet. 1983. Parasitic Gaps. *Linguistics and Philosophy* 6, 5-34.

Evers, Arnold. 1975. *The transformational cycle of Dutch and German*. PhD Dissertation, University of Utrecht.

Fox, Danny and David Pesetsky. 2005. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31: 1-45.

Frank, Robert. 2002. *Phrase Structure Composition and Syntactic Dependencies*. Cambridge, Mass: MIT Press.

Frank, Robert, Young-Suk Lee and Owen Rambow. 1996. Scrambling, Reconstruction and Subject Binding. *Rivista di Grammatica Generativa* 21:67-106.

Geilfuß, Jochen. 1991. Jiddisch als SOV-Sprache. In Arbeitspapiere des Sonderforschungbereichs 340: *Sprachtheoretische Grundlagen für die Computerlinguistik*. Bericht 11: *Verb- und Verbphrasensyntax*, Stuttgart, p. 3-17.

Grosz, Barbara J., Aravind Joshi, and Scott Weinstein. June 1995. Centering: A Framework for Modeling the Local Coherence of Discourse. In *Computational Linguistics* 2(21), pp. 203-225.

Haeberli, Eric. 2002. Inflectional morphology and the loss of verb second in English. In D. Lightfoot ed., *Syntactic Effects of Morphological Change*, 88-106. Oxford: Oxford University Press.

Haeberli, E. 2005. Clause type asymmetries in Old English and the syntax of verb movement. In Batllori, M. and F. Roca eds., *Grammaticalization and Parametric Change*. Oxford: Oxford University Press. Pp. 267-283.

Haeberli, Eric and Liliane Haegeman. 1998 "Negative Concord and Verb Projection Raising in Old English and West Flemish". In I. Tieken-Boon van Ostade, G. Tottie and W. van der Wurff eds., 101-119. *Negation in the History of English*. Berlin/New York: Mouton de Gruyter.

Haegeman, Liliane. 1996. Object Clitics in West Flemish. In Aaron L. Halpern and Arnold M. Zwicky, eds., *Approaching Second: Second Position Clitics and Related Phenomena*, Stanford: CSLI Publications.

Haegeman, Liliane. 1998. V-positions and the middle field in West Flemish. *Syntax* 1:259-299.

Haegeman, Liliane and Henk van Riemsdijk. 1986. Verb Projection Raising, Scope, and the Typology of Rules Affecting Verbs. *Linguistic Inquiry* 17, 3: 417-466.

Haeseryn, Walter. 1990. *Syntactische normen in het Nederlands. Een empirisch onderzoek naar volgordevariatie in de werkwoordelijke eindgroep*. Doctoral Dissertation, Catholic University of Nijmegen.

Haider, Hubert. 1987. Expletives pro – eine Phantomkategorie. In W. U. Dressler, C. Grassi, R. Rindler-Schjerve, and M. Stegu eds., *Parallela* III. Tübingen: Gunter Narr Verlag. pp. 237-247.

Haider, Hubert. 1989. θ-tracking systems – evidence from German. In L. Maracz and P. Muysken eds., *Configurationality: The Typology of Asymmetries*. Dordrecht: Foris. pp. 185-206.

Haider, Hubert. 1993. *Deutsche Syntax generativ – Vorstudien zur Theorie einer projektiven Grammatik*. Tübingen: Gunter Narr Verlag.

Haider, Hubert and Martin Prinzhorn. 1986. *Verb second phenomena in Germanic languages*. Dordrecht: Foris.

Hall, Beatrice. 1979. Accounting for Yiddish Word Order, or What's a nice NP Like You Doing in a Place Like This?. In Jürgen M. Meisel and Martin D. Pam eds., *Linear Order and Generative Theory*. Amsterdam: John Benjamins, p.263-287.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale and Samuel Jay Keyser eds., *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press. 111–176.

Harley, Heidi. 2002. Possession and the double object construction. *Yearbook of Linguistic Variation* 2: 29-68.

Haverkort, Marco. 1993. Romance and Germanic clitics: A comparison of their syntactic behavior. In T. J. Shannon and J. P. Snapper eds., *The Berkeley Conference on Dutch Linguistics 1993*.

Hallgrímur Helgason. 2008. *10 ráð til að hætta að drepa fólk og byrja að vaska upp*. Rekjavík: JPV útgáfa.

Hellan, Lars, and Christer Platzack. 1995. Pronouns in Scandinavian Languages: An Overview. *Working Papers in Scandinavian Syntax* 56: 47-69.

Henry, Alison. 1995. *Belfast English and standard English: dialect variation and parameter setting*. Oxford: Oxford University Press.

Hinterhölzl, Roland. 2000. Licensing Movement and Stranding in the West Germanic OV Languages. In Peter Svenonius ed., *The Derivation of OV and VO*. Linguistik Aktuell/Linguistics Today 31, pp. 293-327. Amsterdam: John Benjamins.

Hinterhölzl, Roland. 2006. *Scrambling, Remnant Movement, and Restructuring in West Germanic*. (Oxford Studies in Comparative Syntax). Oxford: OUP.

Hinterhölzl, Roland. 2009. Information structure, word order and word order change. In Hinterhölzl, Roland and Svetlana Petrova eds., *Information Structure and Language Change*. Berlin: Mouton de Gruyter. 45-66.

Hoekstra, Teun. 1988. Small clause results. *Lingua* 74: 101-139.

Holmberg, Anders. 1984. On Certain Clitic-Like Elements in Swedish. *Working Papers in Scandinavian Syntax* 13: 1-31.

Holmberg, Anders. 1986. *Word order and syntactic features in the Scandinavian languages and English*. Ph.D. diss., University of Stockholm.

Holmberg, Anders. 1997. The true nature of Holmberg's generalization. *NELS* 27, 203-217.

Holmberg, Anders. 1999. Remarks on Holmberg's Generalization. *Studia Linguistica* 53, 1: 1-39.

Holmberg, Anders and Christer Platzack. 1995. *The role of inflection in Scandinavian syntax*. New York: Oxford University Press.

Huygbregts, Riny and Henk van Riemsdijk. 1985. Parasitic Gaps and ATB. In *Proceedings of NELS 15*: 168-187. GLSA, University of Massachusetts, Amherst.

Iatridou, Sabine and Anthony Kroch. 1992. The Licensing of CP-recursion and its Relevance to the Germanic Verb-Second Phenomenon. *Working Papers in Scandinavian Syntax* 50: 1-24.

Jacobson, Pauline, and Paul Neubauer. 1976. Rule cyclicity: Evidence from the intervention constraint. *Linguistic Inquiry*, *7*, 429–462.

Jóhannes Gísli Jónsson. 1996. Clausal Architecture and Case in Icelandic. Ph.D. dissertation, University of Massachusetts at Amherst.

Josefsson, Gunlög. 1992. Object Shift and Weak Pronominals in Swedish. *Working Papers in Scandinavian Syntax* 49: 59-94.

Joshi, Aravind K. and Scott Weinstein. 1981. Control of inference: role of some aspects of discourse structure—Centering. In *Proc. International Joint Conference on Artificial Intelligence*. Vancouver, B.C. 385-87.

Joshi, Aravind K., L. S. Levy and M. Takahashi. 1975. Tree Adjunct Grammars. *Journal of Computer and System Science*, 10:136–163.

Kathol, Andreas. 2001. On the Nonexistence of True Parasitic Gaps in Standard German. In Peter W. Culicover and Paul M. Postal eds., *Parasitic Gaps*. Cambridge: MIT Press.

Kayne, Richard S. 1991. Romance Clitics, Verb Movement, and PRO. *Linguistics Inquiry* 22: 647-686.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Koopman, Hilda. 1984. *The Syntax of Verbs: From Verb Movement Rules in the Kru Language to Universal Grammar*. Dordrecht: Foris.

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

Kiparsky, Paul. 1996. The Shift to Head-Initial VP in Germanic. In H. Thrainsson, J. Peter, and S. Epstein (eds.), Comparative Germanic Syntax. Dordrecht: Kluwer.

Klavans, Judirth L. 1985. The independence of syntax and phonology in cliticization. *Language* 61: 95-120.

Koopman, Hilda. 1984. *The Syntax of Verbs: From Verb Movement Rules in the Kru Language to Universal Grammar*. Dordrecht: Foris.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck and Laurie Zaring eds., *Phrase Structure and the Lexicon*. Dordrecht: Kluwer.

Kroch, A. and A. K. Joshi. 1985. The Linguistic Relevance of Tree Adjoining Grammar. University of Pennsylvania Department of Computer and Information Sciences

Technical Report MS-CIS-85-16.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1: 199-244.

Kroch, Anthony. 1994. Morphosyntactic variation. In K. Beals et al eds., *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*.

Kroch, Anthony and Chunghye Han. 2000. The rise of 'do'-support in English: implications for clause structure. In The Proceedings of the 30th Meeting of the Northeast Linguistics Society (NELS). Rutgers University.

Kroch, Anthony and Susan Pintzuk. 1989. The rightward movement of complements and adjuncts in the Old English of *Beowulf*. *Language Variation and Change* 1: 115-143.

Kroch, Anthony and Beatrice Santorini. 1991. The derived constituent structure of the West Germanic verb-raising construction. In Robert Freidin ed., *Principles and Parameters in Comparative Grammar*, Cambridge: MIT Press.

Kroch, Anthony and Beatrice Santorini. *forthcoming*. Penn Parsed Corpus of Modern British English.

Kroch, Anthony, Beatrice Santorini, and Lauren Delfs. 2004. Penn-Helsinki Parsed Corpus of Early Modern English. http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-1/

Kroch, Anthony and Ann Taylor. 1995. Verb Movement in Old and Middle English: Dialect Variation and Language Contact. In Ans van Kemenade and Nigel Vincent

(eds.), Parameters of Morphosyntactic Change. Cambridge: Cambridge University Press.

Kroch, Anthony & Ann Taylor. 2000. Verb-complement order in Middle English. In S. Pintzuk, G. Tsoulas, and A. Warner (Eds.), Diachronic Syntax: Models and Mechanisms, 132-63. Oxford: Oxford University Press.

Kroch, Anthony, and Ann Taylor. 2000b. Penn-Helsinki Parsed Corpus of Middle English, second edition. http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-2/

Labov, William. 1977. On the use of the present to explain the past. In Adam Makkai, Valerie Becker Makkai, Luigi Heilmann, eds., *Linguistics at the Crossroads*, Padova: Tipografia "La Garangola".

Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

Labov, William. 1994. *Principles of Linguistic Change, Internal Factors*. Oxford: Blackwell.

Larson, Richard K. 1988. On the Double Object Construction. *Linguistic Inquiry* 19, 3: 335-391.

Lebeaux, David. 1995. Where does binding theory apply?. In *University of Maryland Working Papers in Linguistics* 3: 63-88.

Lee, Young-Suk. 1993. *Scrambling as Case-Driven Obligatory Movement*. PhD Dissertation, University of Pennsylvania.

Lee, Young-suk and Beatrice Santorini. 1994. Towards resolving Webelhuth's paradox: Evidence from German and Korean. In Norbert Corver and Henk van Riemsdijk, eds., *Studies on scrambling. Movement and non-movement approaches to free word-order phenomena* (Studies in generative grammar 41). Berlin: Mouton de Gruyter. 257-300.

Lenerz, Jürgen. 1977. Zur Abfolge nominaler Satzglieder im Deutschen. TBL-Verlag Narr.

Light. Caitlin and Joel C. Wallenberg. 2008. The Loss of Quantifier Movement and OV in English: an illusory connection. Paper presented at DIGS 10, August 7, 2008, at Cornell University, Ithaca, N.Y.

Mahajan, Anoop. 1990. *The A/A' Distinction and Movement Theory*. PhD Dissertation, MIT.

Maling, Joan and Annie Zaenen. 1990. Unaccusative, Passive, and Quirky Case. In Joan Maling and Annie Zaenen eds., *Modern Icelandic Syntax*, Syntax and Semantics vol. 24, San Diego: Academic Press.

Marantz, Alec. 1993. Implications of Asymmetries in Double Object Constructions. In Sam A. Mchombo ed., *Theoretical Aspects of Bantu Grammar*. Stanford: CSLI Publications.

McCloskey, James. 1996. On the Scope of Verb-Movement in Irish. *Natural Language and Linguistic Theory* 14: 47-104.

McCloskey, James. 2000. Quantifier Float and *Wh*-Movement in an Irish English. *Linguistic Inquiry* 31, 1: 57-84.

McCloskey, James. 2001. On the Morphosyntax of Wh-Movement in Irish. *Journal of Linguistics* 37.1: 67-100.

Nowak, Martin A. 2006. *Evolutionary Dynamics: exploring the equations of life*. Cambridge, MA: Belknap Press of Harvard University Press.

Olsvanger, Immanuel. 1947. *Royte Pomerantsen*. New York: Schocken.

Pancheva, Roumyana. 2005. The rise and fall of second-position clitics. *Natural Language and Linguistic Theory* 23: 103-167.

Penner, Zvi. 1990. On the Acquisition of Verb Placement and Verb Projection Raising in Bernese Swiss German. In M. Rothweiler ed., *Spracherwerb und Grammatik: linguistische Untersuchungen zum Erwerb von Syntax und Morphologie*. Opladen: Westdeutcher Verlag. 166-189.

Pintzuk, Susan. 1991. *Phrase Structures in Competition: variation and change in Old English word order*. PhD dissertation, University of Pennsylvania.

Pintzuk, Susan. 1996. Cliticization in Old English. In Aaron L. Halpern and Arnold M. Zwicky, eds., *Approaching Second: Second Position Clitics and Related Phenomena*, Stanford: CSLI Publications.

Pintzuk, Susan. 2002. The Syntax of Objects in Old English. Paper given at *Diachronic Issues in Generative Syntax* (DIGS) 7, Girona, Spain.

Pintzuk, Susan. 2005. Arguments against a universal base: evidence from Old English. In English Language *and Linguistics* 9.1, pp. 115-138. Cambridge University Press.

Pintzuk, Susan and Eric Haeberli. 2006. Head-final structure in Old English root clauses. *Diachronic Issues in Generative Syntax* (DIGS) 9, University of Trieste, Italy.

Pintzuk, Susan and Ann Taylor. 2004. The loss of OV order in the history of English. In A. van Kemenade and B. Los (Eds.), *Blackwell Handbook of the History of English*. Oxford: Blackwell.

Pollard, Carl and Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20:33, 365-424.

Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In Cole, Peter, ed. *Radical Pragmatics*. New York: Academic Press. 223-56.

Prince, Ellen F. 1985. Fancy syntax and 'shared knowledge'. *Journal of Pragmatics* 9: 65-81.

Prince, Ellen F. 1999. How not to mark topics: 'Topicalization' in English and Yiddish. In *Texas Linguistics Forum,* Chapt. 8. Austin: University of Texas.

Prince, Ellen F. 2001. Yiddish as a contact language. In Smith, N. and Veenstra, T. eds. *Creolization and contact*. CLL 23. Philadelphia/Amsterdam: John Benjamins, 263-90.

Riemsdijk, H. van. 1985. On pied piped infinitives in German relative clauses. In J. Toman ed., *Issues in the Grammar of German*. Dordrecht: Foris.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Liliane Haegeman ed., *Elements of Grammar*. Dordrecht: Kluwer. 281-337.

Rizzi, Luigi. 2001. Relativized Minimality Effects. In Mark Baltin and Chris Collins eds., *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell. 89-110.

Robbers, Karin. 1997. *Non-finite verbal complements in Afrikaans: A comparative approach*. Ph.D. Dissertation, University of Amsterdam. [Distributed by: Published by Holland Academic Graphics (HIL), Den Haag.]

Roberts, Ian. 1994. Two types of head movement in Romance. in David Lightfoot and Norbert Hornstein eds., *Verb Movement*. Cambridge: Cambridge University Press. 207-242.

Roberts, Ian. 2001. Head Movement. Chapt. 5 in Mark Baltin and Chris Collins eds., *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell.

Roberts, Ian. 2005. *Princicples and Parameters in a VSO Language: a case study in Welsh*. Oxford: OUP.

Robinson, Orrin W. 1997. *Clause Subordination and Verb Placement in the Old High German Isador Translation*. Heidelberg: Universitätsverlag C. Winter.

Rögnvaldsson, Eiríkur. 1982. *Um orðaröð og færslur í íslensku*. Master's thesis, University of Iceland. Published 1990 as vol. 2 of *Málfræðirannsóknir, Málvisinastofnun Háskóla Íslands, Reykjavík*.

Rögnvaldsson, Eiríkur. 1984a. Icelandic Word Order and það-Insertion. *Working Papers in Scandinavian Syntax* 8:1-21.

Rögnvaldsson, Eiríkur. 1984b. Rightward Displacement of NPs in Icelandic. - Formal and Functional Characteristics. In K. Ringgaard & V. Sørensen (eds.), *The Nordic Languages and Modern Linguistics* 5:361-368.

Rögnvaldsson, Eiríkur. 1987. OV Word Order in Icelandic. In R.D.S. Allan and M.P. Barnes eds., *Proceedings of the Seventh Biennial Conference of Teachers of Scandinavian Studies in Great Britain and Northern Ireland*. 33-49. London: University College London.

Rooth, Mats. 1985. *Association with Focus*. PhD Dissertation, University of Massachusetts, Amherst.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1,1:75-116.

Rooth, Mats. 1997. Focus. In Shalom Lappin ed., *Handbook of Contemporary Semantic Theory*. Oxford: Blackwell.

Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Ph.D. dissertation, MIT. Cambridge, Massachusetts.

Saito, Mamoru. 1989. Scrambling as semantically vacuous A'-movement. In Mark Baltin and Anthony S. Kroch eds., *Alternative conceptions of phrase structure*. 182-200. Chicago: University of Chicago Press.

Saito, Mamoru. 1985. *Some Asymmetries in Japanese and Their Theoretical Consequences*. PhD Dissertation, M.I.T.

Saito, Mamoru. 1992. Long-distance scrambling in Japanese. *Journal of East Asian Linguistics*, 1:69-118.

Saito, Mamoru and Naoki Fukui. 1998. Order in Phrase Structure and Movement. *Linguistic Inquiry* 29, 3: 439-474.

Salman Rubashov. 1929. *Yidishe gvies-edes in di shayles-vetshuves fun anhoyb XVtn biz sof XVIItn yorhundert*. (Yiddish court testimony in the responsa literature of the early 15th to the late 17th century). In E. Tsherikover, ed., *Historishe shriftn fun YIVO* (YIVO historical papers), vol. 1. Warsaw: Kultur Liga.

Santorini, Beatrice. 1992. Variation and change in Yiddish subordinate clause word order. *Natural Language and Linguistic Theory* 10: 595-640

Santorini, Beatrice. 1993a. The rate of phrase structure change in the history of Yiddish. *Language Variation and Change*, 5: 257-283.

Santorini, Beatrice. 1993b. Jiddisch als gemischte OV/VO-Sprache. In Werner Abraham and Josef Bayer eds., *Dialektsyntax*, Opladen: Westdeutscher Verlag.

Santorini, Beatrice. 1997/2008. *Penn Yiddish Corpus*. Contact: beatrice@babel.ling.upenn.edu for details.

Sapp, Christopher. 2009. The Verbal Complex from Middle High German to Modern German. *Diachronic Issues in Generative Syntax* (DIGS) 11, 24th July 2009, Unicamp, Campinas, São Paolo, Brazil.

Schönenberger, Manuela. 2001. *Embedded V-to-C in Child Grammar: The Acquisition of Verb Placement in Swiss German*. Dordrecht: Kluwer.

Sisam, Celia. 1951. The Scribal Tradition of the *Lambeth Homilies*. *The Review of English Studies* 2, 6: 105-113.

Speyer, Augustin. 2005. Topicalization and the Trochaic Requirement. *Penn Working Papers in Linguistics* 10, 2: 243-256.

Speyer, Augustin. 2008. *Topicalization and Clash Avoidance: On the interaction of prosody and syntax in the history of English with a few glimpses at German*. PhD Dissertation, University of Pennsylvania.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19: 425-450.

Svenonius, Peter. 1996. The optionality of particle shift. *Working Papers in Scandinavian Syntax* 57: 47-75.

Szabolcsi, Anna. 1981. The semantics of topic-focus articulation. In Jan Groenendijk, Theo Janssen, and Martin Stokhof eds., *Formal methods in the study of language*. 513-41. Amsterdam: Matematisch Centrum.

Taylor, Ann. 1990. *Clitics and Configurationality in Ancient Greek*. PhD dissertation, University of Pennsylvania.

Taylor, Ann, Anthony Warner, Susan Pintzuk, & Frank Beths. 2003. The York-Toronto-Helsinki Parsed Corpus of Old English Prose. Helsington, York: University of York.

Thráinsson, Höskuldur. 1979. *On Complementation in Icelandic*. New York: Garland Press.

Thráinsson, Höskuldur. 1986. V1, V2, V3 in Icelandic. In Hubert Haider and Martin Prinzhorn eds., *Verb second phenomena in Germanic languages*. Dordrecht: Foris. 169–174.

Thráinsson, Höskuldur. 2001. Object Shift and Scrambling. In Mark Baltin and Chris Collins eds., *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell Publishing Ltd.

Thráinsson, Höskuldur. 2007. *The Syntax of Icelandic*. Cambridge: Cambridge University Press.

Travis, Lisa. 1984. Parameters and Effects of Word Order Variation. PhD Dissertation, University of Washington, Seattle.

Vallduví, Enric. 1992. *The Informational Component*. PhD Dissertataion, University of Pennsylvania.

Vijay-Shanker, K. and Aravind K. Joshi. 1985. Some computational properties of Tree Adjoining Grammars. In *Proceedings of ACL*.

Vikner, Sten. 2001. *Verb Movement Variation in Germanic and Optimality Theory*. Habilitationschrift, University of Tübingen.

Walker, M.A., Joshi, Aravind K., and Prince, Ellen F. eds. 1998. *Centering theory in discourse*. Oxford: Oxford University Press.

Wallenberg, Joel. 2007. English Weak Pronouns and Object Shift. *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 27-29 April, 2007, University of California at Berkeley.

Wallenberg, Joel. 2008b. The Decline of Early English Object Clitics. Paper presented at the annual meeting of the Linguistic Society of America, Jan. 6, 2008, Chicago, IL.

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

Webelhuth, Gert. 1989. *Syntactic saturation phenomena and the modern Germanic langauges*. PhD Dissertation, University of Massachusetts, Amherst.

Whitman, J. 2001. Kayne 1994: p. 143, fn. 3. In G. Alexandrova ed., *The Minimalist Parameter*. Amsterdam: John Benjamins, pp.77-100.

Williams, Edwin. 2003. *Representation Theory*. Cambridge, MA: MIT Press.

Weinryb, Beier. 1937. *A pekl briv in yidish fun 1588, Krake-Prag*. (A collection of Yiddish letters from 1588, Cracow--Prague). E. Tsherikover, ed., 1937, *Historishe shriftn fun YIVO* (YIVO historical papers), vol. 2. Vilna: YIVO.

Wurff, Wim van der. 1998. Objects and verbs in Icelandic and fifteenth century English: a word order parallel and its causes. *Lingua* 109: 237-265.

Wurmbrand, Susi. 2004. West Germanic verb clusters: The empirical domain. In Katalin É. Kiss and Henk van Riemsdijk eds., *Verb clusters: A study of Hungarian, German, and Dutch*, p. 43-85. Amsterdam: John Benjamins.

Wurmbrand, Susi. 2005. Verb clusters, verb raising, and restructuring. In eds., Martin Everaert and Henk van Riemsdijk, *The Blackwell Companion to Syntax, volumes I-V,* 227-341. Oxford: Blackwell.

Xia, Fei. 2001. *Automatic Grammar Generation From Two Different Perspectives*. PhD Dissertation, University of Pennsylvania.

Zaenen, Annie. 1979. Infinitival Complements in Dutch. *In Papers from the Fifteenth Regional Meeting, CLS*. Chicago Linguistic Society, University of Chicago.

Zwart, C. Jan-Wouter. 1993. *Dutch Syntax: A Minimalist Approach*. PhD Dissertation, University of Groningen.

Zwart, C. Jan-Wouter. 1996a. Clitics, Scrambling, and Head Movement in Dutch. In Aaron L. Halpern and Arnold M. Zwicky, eds., *Approaching Second: Second Position Clitics and Related Phenomena*, Stanford: CSLI Publications.

Zwart, C. Jan-Wouter. 1996b. Verb clusters in Continental West Germanic dialects. In *Microparametric syntax and dialect variation*, ed. James Black and Virginia Motapanyane, 229-258. Amsterdam: John Benjamins.

Zwart, C. Jan-Wouter. 1997. *Morphosyntax of verb movement: a minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer.

Zwicky, Arnold M. 1977. On Clitics. Indiana University Linguistics Club Publ. Bloomington: Indiana.